

# Attribute Dependencies for Data with Grades

RADIM BELOHLAVEK, Palacky University, Olomouc  
VILEM VYCHODIL, Palacky University, Olomouc

This paper examines attribute dependencies in data that involve grades, such as a grade to which an object is red or a grade to which two objects are similar. We thus extend the classical agenda by allowing graded, or “fuzzy”, attributes instead of Boolean, yes-or-no attributes in case of attribute implications, and allowing approximate match based on degrees of similarity instead of exact match based on equality in case of functional dependencies. In a sense, we move from bivalence, inherently present in the now-available theories of dependencies, to a more flexible setting that involves grades. Such a shift has far-reaching consequences. We argue that a reasonable theory of dependencies may be developed by making use of mathematical fuzzy logic, a recently developed many-valued logic. Namely, the theory of dependencies is then based on a solid logic calculus the same way the classical dependencies are based on classical logic. For instance, rather than handling degrees of similarity in an ad hoc manner, we consistently treat them as truth values, the same way as *true* (match) and *false* (mismatch) are treated in classical theories. In addition, several notions intuitively embraced in the presence of grades, such as a degree of validity of a particular dependence or a degree of entailment, naturally emerge and receive a conceptually clean treatment in the presented approach. In the paper, we discuss motivations, provide basic notions of syntax and semantics, and develop basic results which include entailment of dependencies, associated closure structures, a logic of dependencies with two versions of completeness theorem, results and algorithms regarding complete non-redundant sets of dependencies, relationship to and a possible reductionist interface to classical dependencies, and relationship to functional dependencies over domains with similarity. We also outline future research topics.

Categories and Subject Descriptors: F.4.1 [Mathematical Logic]: Model theory; H.2.8 [Database Applications]: Data mining; I.2.3 [Deduction and Theorem Proving]: Uncertainty, “fuzzy,” and probabilistic reasoning; I.2.4 [Knowledge Representation Formalisms and Methods]: Relation systems

General Terms: Theory

Additional Key Words and Phrases: attribute dependence, grade, similarity, logic, redundancy, functional dependence

## 1. INTRODUCTION

Attribute dependencies are fundamental for understanding and processing data. In the past, dependencies describing various types of attribute relationships have been studied, particularly in relational databases and data analysis/mining. Arguably, the most important dependencies are those of the form

$$A \Rightarrow B \quad (1)$$

where  $A$  and  $B$  are sets of attributes. They are interpreted in two basic ways—in binary datasets (tables with yes-or-no attributes) describing which objects have which attributes and in relations (tables with general attributes) describing the values of objects for the attributes. In binary datasets,  $A \Rightarrow B$  is considered valid if

$$\text{every object (table row) having all attributes from } A \text{ has all attributes from } B, \quad (2)$$

or, more generally, if a certain percentage (called confidence) of the objects having  $A$  also have  $B$  and another percentage (called support) of objects have all the attributes from  $A \cup B$ . Such dependencies are used in data analysis and are known as attribute implications [Carpineto and Romano 2004; Ganter and Wille 1999; Guigues and Duquenne 1986], see also [Delobel and Casey 1973; Fagin 1977], or association rules when the support and confidence are considered [Agrawal et al. 1993; Hájek et al. 2010; Hájek and Havránek 1978; Hipp et al. 2000; Tan et al. 2005]. In relations,  $A \Rightarrow B$  is considered valid if

$$\begin{aligned} &\text{every two tuples (table rows) with the same values on attributes from } A \\ &\text{have the same values on attributes from } B. \end{aligned} \quad (3)$$

Such dependencies are called functional dependencies and are fundamental to relational databases [Armstrong 1974; Codd 1970; Maier 1983].

A common feature of the two interpretations is a bivalent character of the conditions involved in (2) and (3). The bivalence results from the nature of the data. Namely, a given object either does or does not have a given attribute; two given tuples either do or do not have the same value for a given attribute. It turns out that it is becoming increasingly important for data models to account for fuzziness [Fagin 1999; Fagin 2002], which is inherently present

in human cognition and plays a fundamental role in how people communicate knowledge [Zadeh 1988; Zadeh 1994; Zadeh 2008]. Two points in case are fuzzy (or graded) attributes, such as “green” or “performing well”, and similarity relations. In these and other cases, fuzziness is conveniently represented by grades (degrees, scores) which are usually numbers ranging between 0 and 1. Thus, an object  $x$  may be assigned a grade to which  $x$  is green—the higher the grade, the more green  $x$  is. Likewise, two objects  $x$  and  $y$  may be assigned a degree to which  $x$  and  $y$  are similar. A scale of grades bounded by 0 and 1 thus naturally replaces the two-element set of truth values of classical logic with 0 representing falsity (“attribute does not apply”, “values do not match”) and 1 representing truth (“attribute applies”, “values match”). For data with grades, the ordinary dependencies have limited applicability. Namely, rather than knowing that (full) presence of some attributes implies (full) presence of some other attributes, one is naturally interested in rules that take the grades into account. Such rules are the main subject of the present paper.

In particular, we consider rules saying that presence of attributes  $y_i$  with grades at least  $a_i$  implies (or implies partially) presence of attributes  $z_i$  with grades at least  $b_i$ . Therefore, from rules of the form

$$\{y_1, \dots, y_p\} \Rightarrow \{z_1, \dots, z_q\} \quad (4)$$

we come to rules of the form

$$\{^{a_1}/y_1, \dots, ^{a_p}/y_p\} \Rightarrow \{^{b_1}/z_1, \dots, ^{b_q}/z_q\}, \quad (5)$$

such as

$$\{^{0.5}/\text{unhealthy food}, ^{0.9}/\text{little activity}\} \Rightarrow \{^{0.7}/\text{high cholesterol}\}. \quad (6)$$

From a functional dependence point of view, such rules may be interpreted in tables whose domains are equipped with similarity relations assigning similarity grades to pairs of elements of the domains. In such tables, the rules specify that two tuples with similar values on attributes  $y_1, \dots, y_p$  have similar values on  $z_1, \dots, z_q$ . In particular, rule (5) says that similarity to degrees  $a_i$  or higher on attributes  $y_i$  implies similarity to degrees  $b_i$  or higher on attributes  $z_i$ , generalizing thus ordinary functional dependencies which say that a match of two tuples on attributes  $y_1, \dots, y_p$  implies a match on  $z_1, \dots, z_q$ .

Using grades to represent fuzziness is the fundamental idea of fuzzy logic [Zadeh 1965]. We use fuzzy logic as a formal framework for our approach. Fuzzy logic enables us to manipulate the grades by means of the truth functions of logic connectives. In the past, various models of processing data with grades using fuzzy logic connectives, notably “fuzzy conjunction”, have been studied in a more or less ad hoc way. In this perspective, one aspect of our work is that we consistently use the so-called mathematical fuzzy logic [Cintula et al. 2011; Gottwald 2001; Gottwald 2008; Hájek 1998; Hájek 2006] as a formal framework. Mathematical fuzzy logic is a recently developed branch of logic that provides us with general principles and notions such as theory, model, or entailment, and enables us to process data with grades in a clean way.

Our reliance on mathematical fuzzy logic is similar to the reliance of the ordinary dependencies on classical logic. In case of grades, however, the logic framework is more explicit. Namely, while in the ordinary case the assertions like (2) or the notion of entailment have a clear meaning and one rarely needs to resort to the formal agenda of classical logic, in case of grades, the meaning needs to be supplied by an explicit resorting to fuzzy logic principles. Due to a consistent use of fuzzy logic, the verbal description of validity conditions and manipulation regarding the dependencies remains the same as in the ordinary case, retaining thus a clear meaning. For instance, the validity of rule (5) in data with grades may still be verbally described by (2), the grades being “hidden in the interpretation”. A natural consequence of working with grades is that key logic notions such as validity or entailment become graded. That is, we speak of a degree to which a given rule is valid or a degree to which a rule follows from other rules leaving validity or entailment to degree 1 (full validity or full entailment) important particular cases.

While the reliance on mathematical fuzzy logic provides us with reasonable guiding principles, the resulting notions and problems tend, naturally, to be more involved both conceptually as well as technically due to the presence of intermediary grades and, in addition, due to the fact that we develop the theory for a general scale  $L$  of truth degrees with  $L$  acting as a parameter. The conceptual aspect regarding the extension from the ordinary, bivalent framework to a framework involving grades may, using a loose analogy, be compared to an extension from a deterministic to a probabilistic framework. As regards the technical aspect, a point in case for illustration is the fact that, as a rule, the ordinary proofs by cases, corresponding

to *false* and *true*, no longer work and need to be replaced by different schemata which are based on algebraic manipulation of the grades. In this perspective, the paper illustrates both aspects, the conceptual and the technical, by numerous cases.

The paper is organized as follows. In Section 2, we present preliminaries on scales of truth degrees and operations on them and the basic principles of fuzzy logic. Section 3 presents the basic notions regarding graded attribute implications, their validity, theories, models, entailment, and related closure and other structures. In Section 4, we present a system for reasoning with graded attribute implications that is based on Armstrong-like rules and prove two versions of syntactico-semantical completeness, the ordinary-style one claiming that entailment coincides with provability and the graded-style one claiming that degrees of entailment equal degrees of provability. Section 5 elaborates on the notion of a base, that is a non-redundant set of graded attribute implications that contains, via entailment, complete information about validity of all implications in a given data. In particular, we focus on bases constructed by means of so-called pseudo-intents. The algorithms for the problem of computing bases and some other problems regarding attribute implications are presented in Section 6. In Section 7, we explore the problem of whether and to what extent it is possible to reduce the notions and problems regarding graded attribute implications, notably the problem of computing a base, to the corresponding problems regarding ordinary attribute implications. Section 8 presents the above mentioned alternative semantics for graded attribute implications in which implications are interpreted as functional dependencies over a certain extension of Codd's relational model, in which domains of attribute values are equipped with binary fuzzy relations. The binary relations may, in particular, be preference relations or similarity relations, in which case the extension becomes a relational model enabling similarity queries and other data processing involving similarity relations. We examine such extension in detail in another paper. In this paper, we show that the two semantics are equivalent in that their notions of (degree of) entailment coincide.

## 2. SCALES OF GRADES AND BASIC PRINCIPLES OF FUZZY LOGIC

The dependencies studied in this paper are of the form (5) and we assume that they are interpreted in tables with graded attributes. We assume that the grades involved (i.e.  $a_i$ s,  $b_i$ s, and the table entries) belong to a fixed set  $L$ . Furthermore, we assume that  $L$  is bounded by 0 and 1, partially ordered (usually a chain), and equipped with operations which are (truth functions of) logic connectives. In accordance with fuzzy logic, we interpret the grades in  $L$  as truth values, or truth degrees, with 0 and 1 representing falsity and truth. The intermediate degrees  $a$ , i.e. those with  $0 < a < 1$ , represent partial truth. As in classical logic, grades are assigned to propositions to represent their validity. The grade assigned to proposition  $\varphi$  in structure  $M$  is denoted by

$$||\varphi||_M \text{ or just } ||\varphi||.$$

Higher grades indicate truer propositions, hence

$$||x \text{ is red}|| = 0.7 \quad \text{and} \quad ||y \text{ is red}|| = 0.9$$

implies that  $y$  is considered more red than  $x$ . We consider (truth functions of) conjunction and implication and denote them by  $\otimes$  and  $\rightarrow$ . As usual in fuzzy logic, we assume truth functionality of connectives. That is, the truth degree of  $\varphi \& \psi$  and  $\varphi \Rightarrow \psi$  (conjunction and implication of  $\varphi$  and  $\psi$ ) is defined as

$$||\varphi \& \psi|| = ||\varphi|| \otimes ||\psi|| \quad \text{and} \quad ||\varphi \Rightarrow \psi|| = ||\varphi|| \rightarrow ||\psi||.$$

This way, the operations may be looked at as aggregation operations [Fagin 2002]. For instance, if

$$||x \text{ is brown}|| = 0.8 \quad \text{and} \quad ||x \text{ is heavy}|| = 0.5,$$

and if  $\otimes$  is the Goguen conjunction (see below in this section), then the degree to which  $x$  is brown and heavy is  $0.8 \otimes 0.5 = 0.8 \cdot 0.5 = 0.4$ . To be able to evaluate truth degrees of quantified formulas, we assume that as a partially ordered set,  $L$  forms a complete lattice, i.e. infima and suprema of arbitrary sets of grades exist. Namely, if  $\varphi$  is a formula with a free variable  $x$  ranging over a set  $D$ , one naturally defines

$$||(\forall x)\varphi|| = \bigwedge_e ||\varphi||_e \quad \text{and} \quad ||(\exists x)\varphi|| = \bigvee_e ||\varphi||_e,$$

where  $e$  ranges over all valuations of  $x$  in  $D$ . It has been recognized in his seminal work by Goguen [Goguen 1967; Goguen 1968–9] that a class of general scales of grades

equipped with operations suitable for fuzzy logic is the class of all complete residuated lattices [Ward and Dilworth 1939]. Residuated lattices and their variants are currently the main structures used in mathematical fuzzy logic [Galatos et al. 2007; Gottwald 2001; Gottwald 2008; Hájek 1998] and are used as the basic structures of grades in this paper.

A complete residuated lattice [Hájek 1998; Hájek 2001] is an algebra  $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1 \rangle$  such that  $\langle L, \wedge, \vee, 0, 1 \rangle$  is a complete lattice with 0 and 1 being the least and greatest elements, respectively;  $\langle L, \otimes, 1 \rangle$  is a commutative monoid (i.e.  $\otimes$  is commutative, associative, and  $a \otimes 1 = 1 \otimes a = a$  for each  $a \in L$ ); and  $\otimes$  with  $\rightarrow$  satisfy the so-called adjointness property:

$$a \otimes b \leq c \quad \text{iff} \quad a \leq b \rightarrow c \quad (7)$$

for each  $a, b, c \in L$ . Commonly used residuated lattices are those with  $L = [0, 1]$  (unit interval),  $\wedge$  and  $\vee$  being minimum and maximum,  $\otimes$  being a left-continuous t-norm [Gottwald 2001; Hájek 1998] and  $\rightarrow$  its residuum. Three most important pairs of adjoint operations on the unit interval are:

$$\begin{array}{ll} \text{\textbf{Łukasiewicz:}} & \begin{array}{l} a \otimes b = \max(a + b - 1, 0), \\ a \rightarrow b = \min(1 - a + b, 1), \end{array} \end{array} \quad (8)$$

$$\begin{array}{ll} \text{\textbf{Gödel:}} & \begin{array}{l} a \otimes b = \min(a, b), \\ a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b, \\ b & \text{otherwise,} \end{cases} \end{array} \end{array} \quad (9)$$

$$\begin{array}{ll} \text{\textbf{Goguen (product):}} & \begin{array}{l} a \otimes b = a \cdot b, \\ a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b, \\ \frac{b}{a} & \text{otherwise.} \end{cases} \end{array} \end{array} \quad (10)$$

Another important class of examples consists of residuated lattices that are finite equidistant subchains in  $[0, 1]$ , i.e.  $L = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$ . Such chains may be endowed with the restrictions of Łukasiewicz, Gödel operations, or other discrete t-norm-based operations [Mayor and Torrens 2005]. Importantly, a particular example for  $n = 1$  yields  $L = \{0, 1\}$  in which case  $\otimes$  and  $\rightarrow$  are the classical conjunction and implication. In this case,  $\mathbf{L}$  is the two-element Boolean algebra of classical logic and is denoted by  $\mathbf{2}$  in this paper.

The following are the basic properties of complete residuated lattices that are needed in our paper, see e.g. [Belohlavek 2002; Gottwald 2001; Hájek 1998]:

**THEOREM 2.1.** *Every complete residuated lattice satisfies*

$$a \leq b \quad \text{iff} \quad a \rightarrow b = 1, \quad (11)$$

$$a \rightarrow a = 1, \quad a \rightarrow 1 = 1, \quad 0 \rightarrow a = 1, \quad (12)$$

$$1 \rightarrow a = a, \quad (13)$$

$$a \otimes 0 = 0, \quad (14)$$

$$a \otimes b \leq a, \quad a \leq b \rightarrow a, \quad (15)$$

$$a \otimes (a \rightarrow b) \leq b, \quad b \leq a \rightarrow (a \otimes b), \quad a \leq (a \rightarrow b) \rightarrow b, \quad (16)$$

$$(a \otimes b) \rightarrow c = a \rightarrow (b \rightarrow c) = b \rightarrow (a \rightarrow c), \quad (17)$$

$$(a \rightarrow b) \otimes (b \rightarrow c) \leq a \rightarrow c, \quad (18)$$

$$a_1 \leq a_2 \text{ and } b_1 \leq b_2 \text{ implies } a_1 \otimes b_1 \leq a_2 \otimes b_2, \quad (19)$$

$$a_1 \geq a_2 \text{ and } b_1 \leq b_2 \text{ implies } a_1 \rightarrow b_1 \leq a_2 \rightarrow b_2, \quad (20)$$

$$(a \rightarrow b) \otimes (c \rightarrow d) \leq (a \otimes c) \rightarrow (b \otimes d), \quad (21)$$

$$a \otimes \bigvee_{i \in I} b_i = \bigvee_{i \in I} (a \otimes b_i), \quad (22)$$

$$a \rightarrow \bigwedge_{i \in I} b_i = \bigwedge_{i \in I} (a \rightarrow b_i), \quad (23)$$

$$\bigvee_{i \in I} a_i \rightarrow b = \bigwedge_{i \in I} (a_i \rightarrow b), \quad (24)$$

$$a \otimes \bigwedge_{i \in I} b_i \leq \bigwedge_{i \in I} (a \otimes b_i), \quad (25)$$

$$\bigvee_{i \in I} (a \rightarrow b_i) \leq a \rightarrow \bigvee_{i \in I} b_i, \quad (26)$$

$$\bigvee_{i \in I} (a_i \rightarrow b) \leq \bigwedge_{i \in I} a_i \rightarrow b, \quad (27)$$

$$\bigwedge_{i \in I} (a_i \rightarrow b_i) \leq \bigwedge_{i \in I} a_i \rightarrow \bigwedge_{i \in I} b_i. \quad (28)$$

Residuated lattices may be equipped with further operations. We utilize truth-stressing hedges (shortly, hedges) which are functions  $*$  :  $L \rightarrow L$  that represent intensifying linguistic

modifiers such as “very” or “highly”. Such modifiers are used in propositions like “this book is very good” or, put differently, “it is very true that this book is good”, and may be thought of as unary logic connectives [Hájek 1998; Hájek 2001]. If  $*$  is the hedge representing the modifier “very”, then the truth degree of the proposition “it is very true that  $\varphi$ ”, shortly “very  $\varphi$ ”, is  $\|\varphi\|^*$ . That is, one applies  $*$  to the truth degree of  $\varphi$ . We assume that a truth-stressing hedge satisfies the following conditions, which are inspired by the conditions from [Hájek 2001]:

$$1^* = 1, \quad (29)$$

$$a^* \leq a, \quad (30)$$

$$(a \rightarrow b)^* \leq a^* \rightarrow b^*, \quad (31)$$

$$a^{**} = a^*, \quad (32)$$

for each  $a, b \in L$  ( $i \in I$ ). Properties (29)–(32) have a natural interpretation. For instance, (29) says that if a proposition  $\varphi$  is true (to degree 1), it is also very true (to degree 1). (30) says that if  $\varphi$  is very true, then  $\varphi$  is true; (31), which is equivalent to  $a^* \otimes (a \rightarrow b)^* \leq b^*$ , says that if  $\varphi$  is very true and  $\varphi \Rightarrow \psi$  is very, then  $\psi$  is very true; and (32) says that “very very  $\varphi$ ” has the same truth degree as “very  $\varphi$ ”.

Two boundary cases of hedges are (i) identity, i.e.  $a^* = a$  ( $a \in L$ ); (ii) globalization [Takeuti and Titani 1987]:

$$a^* = \begin{cases} 1 & \text{if } a = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (33)$$

Note that identity is the only hedge on the two-element Boolean algebra **2**.

Given a complete residuate lattice **L**, one defines the usual notions regarding fuzzy sets: an **L**-set (fuzzy set)  $A$  in universe  $U$  is a mapping  $A: U \rightarrow L$ ,  $A(u)$  being interpreted as “the degree to which  $u$  belongs to  $A$ ”. If  $U = \{u_1, \dots, u_n\}$  then  $A$  can be denoted by  $A = \{a_1/u_1, \dots, a_n/u_n\}$  meaning that  $A(u_i)$  equals  $a_i$  for each  $i = 1, \dots, n$ . For brevity, we introduce the following convention: we write  $\{\dots, u, \dots\}$  instead of  $\{\dots, 1/u, \dots\}$ , and we also omit elements of  $U$  whose membership degree is zero. For example, we write  $\{u, 0.5/v\}$  instead of  $\{1/u, 0.5/v, 0/w\}$ , etc. Let  $L^U$  or  $\mathbf{L}^U$  (if the operations on  $L$  are to be emphasized) denote the collection of all **L**-sets in  $U$ . The basic operations with **L**-sets are defined componentwise. For instance, the intersection of **L**-sets  $A, B \in \mathbf{L}^U$  is an **L**-set  $A \cap B$  in  $U$  such that  $(A \cap B)(u) = A(u) \wedge B(u)$  for each  $u \in U$ , etc. Binary **L**-relations (binary fuzzy relations) between  $X$  and  $Y$  can be thought of as **L**-sets in the universe  $X \times Y$ . That is, a binary **L**-relation  $I \in \mathbf{L}^{X \times Y}$  between a set  $X$  and a set  $Y$  is a mapping assigning to each  $x \in X$  and each  $y \in Y$  a truth degree  $I(x, y) \in L$  (a degree to which  $x$  and  $y$  are related by  $I$ ). An **L**-set  $A \in \mathbf{L}^X$  is called *crisp* if  $A(x) \in \{0, 1\}$  for each  $x \in X$ . Crisp **L**-sets may obviously be identified with ordinary sets. For a crisp  $A$ , we also write  $x \in A$  if  $A(x) = 1$  and  $x \notin A$  if  $A(x) = 0$ . An **L**-set  $A \in \mathbf{L}^X$  is called *empty* (denoted by  $\emptyset$ ) if  $A(x) = 0$  for each  $x \in X$ .

For  $a \in L$  and  $A \in \mathbf{L}^X$ , the **L**-sets  $a \otimes A \in \mathbf{L}^X$  and  $a \rightarrow A \in \mathbf{L}^X$  are defined by  $(a \otimes A)(x) = a \otimes A(x)$  and  $(a \rightarrow A)(x) = a \rightarrow A(x)$ .

### 3. GRADED ATTRIBUTE IMPLICATIONS AND THEIR SEMANTICS

#### 3.1. Definition and Validity in Tables with Grades

Throughout the paper, we assume that  $Y$  is a finite and nonempty set of attributes. The dependencies we consider, such as (6), are defined as follows.

*Definition 3.1.* A (graded) attribute implication over  $Y$  is an expression  $A \Rightarrow B$ , where  $A, B \in \mathbf{L}^Y$  ( $A$  and  $B$  are **L**-sets of attributes in  $Y$ ).

Note that since both  $A$  and  $B$  may be crisp in  $A \Rightarrow B$ , i.e.  $A(y), B(y) \in \{0, 1\}$  for each  $y \in Y$ , ordinary attribute implications (association rules, functional dependencies) are a particular case of graded attribute implications. In addition, if **L** is the two-element Boolean algebra, graded implications become just the ordinary attribute implications.

Graded attribute implications are to be interpreted in tables whose entries contain grades to which objects (represented by rows) have attributes (represented by columns). Such tables are represented as triplets  $\langle X, Y, I \rangle$  consisting of non-empty sets  $X$  of objects and  $Y$  of attributes and an **L**-relation  $I: X \times Y \rightarrow L$  for which the degree  $I(x, y)$  is interpreted as the grade to which the attribute  $y \in Y$  applies to the object  $x \in X$ .

Consider first the implication

$$\{1/y_1, 0.5/y_3\} \Rightarrow \{0.8/y_2, 1/y_4\} \quad (34)$$

and the table

$I$	$y_1$	$y_2$	$y_3$	$y_4$
$x_1$	1.0	0.9	0.8	1.0
$x_2$	1.0	0.7	0.8	1.0
$x_3$	0.9	0.5	0.8	1.0

(35)

On intuitive grounds, (34) is satisfied by the object  $x_1$  because  $x_1$  has all the attributes from the antecedent  $A = \{^1/y_1, ^{0.5}/y_3\}$  to the specified grades, i.e.  $A(y_1) \leq I(x_1, y_1)$  and  $A(y_3) \leq I(x_1, y_3)$ , and has also the attributes from the consequent  $B = \{^{0.8}/y_2, ^1/y_4\}$  to the specified grades, since  $B(y_2) \leq I(x_1, y_2)$  and  $B(y_4) \leq I(x_1, y_4)$ . While  $x_2$  has the objects from  $A$  to the specified grades as well,  $y_2$  applies to  $x_2$  to grade 0.7 which is smaller than the grade 0.8 prescribed by  $B$ . Since 0.7 is only slightly smaller than 0.8, one naturally considers (34) as an implication which is *almost satisfied* by the object  $x_2$ , that is, satisfied to a high degree. The object  $x_3$  does not have the attributes from  $A$  to the specified grades, because it possesses the attribute  $y_1$  to grade 0.9 while the grade prescribed by  $A$  is 1. In testing the validity of (34) in the table, one may therefore want to disregard  $x_3$ . However, if one wishes to work consistently with partial satisfiability, the same way one works with classic satisfiability, one should involve  $x_3$  and modify the test to take into account that  $x_3$  satisfies the antecedent  $A$  partially. Clearly, both approaches, one in which only the objects fully satisfying  $A$  participate in testing the validity of  $A \Rightarrow B$  and the other in which also objects partially satisfying  $A$  participate in the test, coincide in the classical case with 0 and 1 as the only grades. In the general case with intermediate grades involved, both approaches are plausible and lead to two, different kinds of semantics. As we show next, it turns out that both of the approaches can conveniently be regarded as two particular cases of a general way to assess validity of  $A \Rightarrow B$  that is parameterized by how one evaluates the satisfaction of  $A$ .

We now provide a definition of validity of a graded attribute implication  $A \Rightarrow B$  in a table  $\langle X, Y, I \rangle$  with grades. The basic structures in which  $A \Rightarrow B$  is evaluated are L-sets of attributes. The rationale is that every row of  $\langle X, Y, I \rangle$  corresponding to the object  $x \in X$  may be seen as the L-set  $I_x \in \mathbf{L}^Y$  given by  $I_x(y) = I(x, y)$  for every  $y \in Y$ . Consider thus  $M \in \mathbf{L}^Y$  representing object  $x$ , i.e.  $M(y)$  is interpreted as the grade to which  $x$  has  $y$ . According to (2), the truth degree  $\|A \Rightarrow B\|_M$  to which  $A \Rightarrow B$  is valid in  $M$  is intended to be the truth degree of the proposition “if  $x$  has all attributes from  $A$  then  $x$  has all attributes from  $B$ ”, or equivalently, “if  $A$  is contained in  $M$  then  $B$  is contained in  $M$ ”.

Containment of an L-set  $C$  in an L-set  $D$  is conveniently represented by the degree  $S(C, D)$  of inclusion of  $C$  in  $D$  [Gottwald 2001], defined by

$$S(C, D) = \bigwedge_{y \in Y} (C(y) \rightarrow D(y)). \quad (36)$$

$S(C, D)$  is the truth degree of proposition “for each  $y \in Y$ : if  $y$  belongs to  $C$  then  $y$  belongs to  $D$ ”. Clearly,  $S$  is a graded relation which generalizes the inclusion relation of classical sets in that if  $L = \{0, 1\}$  then  $S$  is just the characteristic function of classical inclusion. In particular, we write  $C \subseteq D$  if  $S(C, D) = 1$ . As a consequence of the fact that  $a \rightarrow b = 1$  iff  $a \leq b$  we get that  $C \subseteq D$  if and only if  $C(y) \leq D(y)$  for each  $y \in Y$ , i.e.  $C \subseteq D$  means that  $C$  is “fully contained” in  $D$ . In what follows we use the well-known properties of graded inclusion [Gottwald 2001].

With  $S(A, M)$  and  $S(B, M)$  being the degrees to which  $A$  and  $B$  are contained in  $M$ , respectively, one can define the degree to which  $A \Rightarrow B$  is valid in  $M$  by  $\|A \Rightarrow B\|_M = S(A, M) \rightarrow S(B, M)$ . We provide a slightly more general definition to account for both approaches described above, utilizing the notion of hedge introduced in Section 2.

**Definition 3.2.** Let  $\mathbf{L}$  be a complete residuated lattice  $\mathbf{L}$  with a truth-stressing hedge  $*$ . The degree  $\|A \Rightarrow B\|_M \in L$  to which  $A \Rightarrow B$  is valid in an L-set  $M$  of attributes is defined by

$$\|A \Rightarrow B\|_M = S(A, M)^* \rightarrow S(B, M). \quad (37)$$

**Remark 3.3.** (a) If  $*$  is the globalization, i.e.  $a^* = 1$  for  $a = 1$  and  $a^* = 0$  for  $a < 1$ , we get  $\|A \Rightarrow B\|_M = S(B, M)$  if  $A \subseteq M$  and  $\|A \Rightarrow B\|_M = 1$  if  $A \not\subseteq M$ . Namely, if  $A \subseteq M$  then  $S(A, M)^* \rightarrow S(B, M) = 1^* \rightarrow S(B, M) = 1 \rightarrow S(B, M) = 1$ , and if  $A \not\subseteq M$ , i.e.  $S(A, M) < 1$ , then  $S(A, M)^* \rightarrow S(B, M) = 0 \rightarrow S(B, M) = 1$ . This corresponds to the first approach mentioned above, in which only objects fully satisfying  $A$  participate in testing validity. In addition,  $A \Rightarrow B$  is fully true, i.e.  $\|A \Rightarrow B\|_M = 1$ , if and only if

$$A \subseteq M \text{ implies } B \subseteq M. \quad (38)$$

In this case, the degrees  $A(y)$  and  $B(y)$  may be seen as thresholds. Namely, according to (38),  $A \Rightarrow B$  is satisfied by the object  $x$  represented by  $M$  if and only if each attribute  $y \in Y$  applies to  $x$  in grade at least  $A(y)$ , then each attribute  $y \in Y$  applies to  $x$  in grade at least  $B(y)$ .

(b) If  $*$  is the identity, then  $\|A \Rightarrow B\|_M = S(A, M) \rightarrow S(B, M)$ . This corresponds to the second approach mentioned above, in which also objects partially satisfying  $A$  participate in the test of validity. In addition, since  $a \rightarrow b = 1$  iff  $a \leq b$  for any  $a, b \in L$ ,  $A \Rightarrow B$  is fully true if and only if

$$S(A, M) \leq S(B, M). \quad (39)$$

(c) Globalization and identity represent the two natural ways to interpret graded attribute implications. In what follows, we develop the results for general hedges  $*$ , covering thus both globalization and identity as particular cases.

For a system  $\mathcal{M}$  of  $L$ -sets in  $Y$ , the degree  $\|A \Rightarrow B\|_{\mathcal{M}}$  to which  $A \Rightarrow B$  is valid in (each  $M$  from)  $\mathcal{M}$  is defined by

$$\|A \Rightarrow B\|_{\mathcal{M}} = \bigwedge_{M \in \mathcal{M}} \|A \Rightarrow B\|_M. \quad (40)$$

The degree  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  to which  $A \Rightarrow B$  is valid in a table  $\langle X, Y, I \rangle$  with grades is defined by

$$\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \|A \Rightarrow B\|_{\{I_x \mid x \in X\}}. \quad (41)$$

Recall that  $I_x$  represents the  $x$ th row in  $\langle X, Y, I \rangle$ , i.e. for each  $y \in Y$ ,

$$I_x(y) = I(x, y). \quad (42)$$

Hence  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  is naturally interpreted as the degree to which  $A \Rightarrow B$  is valid in every row of table  $\langle X, Y, I \rangle$ .

*Example 3.4.* Consider again the implication  $A \Rightarrow B$  in (34), the table in (35), and the informal requirements discussed above in this example. Let  $L$  be the complete residuated lattice given by the Łukasiewicz operations on  $[0, 1]$ . Since  $S(A, I_{x_1}) = 1$  and  $S(B, I_{x_1}) = 1$ , we get

$$\|A \Rightarrow B\|_{I_{x_1}} = S(A, I_{x_1})^* \rightarrow S(B, I_{x_1}) = 1 \rightarrow 1 = 1.$$

$A \Rightarrow B$  is thus fully satisfied by  $x_1$ , independently of the choice of  $*$  because  $1^* = 1$  is always the case. For  $x_2$ , we have again  $S(A, I_{x_2}) = 1$  but in this case,  $S(B, I_{x_2}) = \bigwedge_{y \in Y} (B(y) \rightarrow I(x_2, y)) = B(y_2) \rightarrow I(x_2, y_2) = 0.8 \rightarrow 0.7 = 0.9$ , whence

$$\|A \Rightarrow B\|_{I_{x_2}} = S(A, I_{x_2})^* \rightarrow S(B, I_{x_2}) = 1 \rightarrow 0.9 = 0.9,$$

again independently of the choice of  $*$ . This corresponds to the intuitive requirement that  $A \Rightarrow B$  be almost satisfied by  $x_2$  because the grades specified by  $B$  are almost attained by the object  $x_2$ . For  $x_3$ , we have  $S(A, I_{x_3}) = A(y_1) \rightarrow I(x_3, y_1) = 1 \rightarrow 0.9 = 0.9$ , i.e.  $A$  is only partially satisfied by  $x_3$ . According to the first approach to the semantics of implications,  $x_3$  should not participate in the test of validity. Indeed, for  $*$  being globalization which corresponds to the first approach, we obtain

$$\|A \Rightarrow B\|_{I_{x_3}} = S(A, I_{x_3})^* \rightarrow S(B, I_{x_3}) = 0 \rightarrow S(B, I_{x_3}) = 1,$$

because  $0 \rightarrow a = 1$  for any degree  $a$ . For  $*$  being the identity, we get

$$\|A \Rightarrow B\|_{I_{x_3}} = S(A, I_{x_3}) \rightarrow S(B, I_{x_3}) = 0.9 \rightarrow 0.7 = 0.8,$$

which corresponds to the second approach to the semantics. We see that  $x_3$  enters the test of validity in that the degree  $S(B, I_{x_3}) = 0.7$  to which  $x_3$  satisfies  $B$  is modified by the degree  $S(A, I_{x_3}) = 0.9$  to which  $x_3$  satisfies  $A$ . In particular, the modification is accomplished by shifting up the degree  $S(B, I_{x_3})$ ; the smaller  $S(A, I_{x_3})$  the more significant the shift. This is because we always have  $S(A, I_{x_3}) \rightarrow S(B, I_{x_3}) \geq S(B, I_{x_3})$  and because  $\rightarrow$  is antitone in the first argument.

This example also makes it clear that testing to what degree an object  $x$  satisfies a consequent  $B$  (or antecedent  $A$ ) actually amounts to comparing the degrees  $B(y)$  and  $I(x, y)$  for every attribute  $y$ . If  $B(y) \leq I(x, y)$ , the test is passed with degree 1 for attribute  $y$ . If  $B(y) \not\leq I(x, y)$ , the test is passed with degree  $B(y) \rightarrow I(x, y) < 1$  for attribute  $y$ . In the end, the thus obtained degrees are aggregated by means of infimum which yields the degree to which  $x$  satisfies  $B$ . Taking  $B(y) \rightarrow I(x, y)$  if  $B(y) \not\leq I(x, y)$  is appropriate because  $\rightarrow$  is antitone in the first and isotone in the second argument. For example, for the Łukasiewicz operations,  $B(y) \rightarrow I(x, y)$  is  $1 - B(y) + I(x, y)$ ; for the Goguen operations,  $B(y) \rightarrow I(x, y) = I(x, y)/B(y)$ .

### 3.2. Theories, Models, and Semantic Entailment

When reasoning with degrees, theories are naturally conceived as  $\mathbf{L}$ -sets of formulas. A (*graded*) *theory* is therefore an  $\mathbf{L}$ -set of graded implications over  $Y$ . The degree  $T(A \Rightarrow B)$  is considered as the degree to which we assume the validity of  $A \Rightarrow B$ . This approach corresponds to the notion of a theory as an  $\mathbf{L}$ -set (fuzzy set) of axioms in fuzzy logic [Pavelka 1979], see also [Gerla 2001; Hájek 1998; Novák et al. 1999]. If  $T$  is a *crisp theory*, i.e.  $T(A \Rightarrow B)$  is 0 or 1 for every  $A \Rightarrow B$ , we write  $A \Rightarrow B \in T$  if  $T(A \Rightarrow B) = 1$  and  $A \Rightarrow B \notin T$  if  $T(A \Rightarrow B) = 0$ .

For a theory  $T$ , the set  $\text{Mod}(T)$  of all *models* of  $T$  is defined by

$$\text{Mod}(T) = \{M \in \mathbf{L}^Y \mid \text{for each } A, B \in \mathbf{L}^Y : T(A \Rightarrow B) \leq \|A \Rightarrow B\|_M\}.$$

That is,  $M \in \text{Mod}(T)$  means that for each attribute implication  $A \Rightarrow B$ , the degree to which  $A \Rightarrow B$  holds in  $M$  is higher than or at least equal to the degree  $T(A \Rightarrow B)$  prescribed by  $T$ . In particular, for a crisp  $T$  we have  $\text{Mod}(T) = \{M \in \mathbf{L}^Y \mid \text{for each } A \Rightarrow B \in T : \|A \Rightarrow B\|_M = 1\}$ .

The degree  $\|A \Rightarrow B\|_T \in L$  to which  $A \Rightarrow B$  *semantically follows* from a fuzzy set  $T$  of attribute implications is defined by

$$\|A \Rightarrow B\|_T = \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow B\|_M. \quad (43)$$

That is,  $\|A \Rightarrow B\|_T$  may be seen as the degree to which  $A \Rightarrow B$  is valid in every model of  $T$ .

We need the following lemma.

- LEMMA 3.5.** (i):  $c \rightarrow S(B, M) = S(c \otimes B, M) = S(B, c \rightarrow M)$ ;  
(ii):  $c \rightarrow \|A \Rightarrow B\|_M = \|A \Rightarrow c \otimes B\|_M$ .  
(iii):  $c \leq \|A \Rightarrow B\|_M$  iff  $\|A \Rightarrow c \otimes B\|_M = 1$ .

**PROOF.** (i): On account of (23) and (17) and we have

$$\begin{aligned} c \rightarrow S(B, M) &= c \rightarrow \bigwedge_{y \in Y} (B(y) \rightarrow M(y)) = \bigwedge_{y \in Y} (c \rightarrow (B(y) \rightarrow M(y))) = \\ &= \bigwedge_{y \in Y} ((c \otimes B(y)) \rightarrow M(y)) = \bigwedge_{y \in Y} ((c \otimes B)(y) \rightarrow M(y)) = S(c \otimes B, M). \end{aligned}$$

$S(c \otimes B, M) = S(B, c \rightarrow M)$  is an easy consequence of (17).

(ii): Using (17) and (i),

$$\begin{aligned} c \rightarrow \|A \Rightarrow B\|_M &= c \rightarrow (S(A, M)^* \rightarrow S(B, M)) = S(A, M)^* \rightarrow (c \rightarrow S(B, M)) = \\ &= S(A, M)^* \rightarrow S(c \otimes B, M) = \|A \Rightarrow c \otimes B\|_M. \end{aligned}$$

(iii): Direct consequence of (ii) and (11).  $\square$

Lemma 3.5 implies every graded theory may be transformed to a crisp theory with the same models and thus (degrees of) consequences:

**THEOREM 3.6.** *Let  $T$  be a theory,  $A \Rightarrow B$  be a graded attribute implication. For the crisp theory  $\text{cr}(T)$  defined by*

$$\text{cr}(T) = \{A \Rightarrow T(A \Rightarrow B) \otimes B \mid A, B \in \mathbf{L}^Y \text{ and } T(A \Rightarrow B) \otimes B \neq \emptyset\} \quad (44)$$

*we have*

$$\text{Mod}(T) = \text{Mod}(\text{cr}(T)), \quad (45)$$

$$\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{\text{cr}(T)}. \quad (46)$$

**PROOF.** (45) directly using (iii) of Lemma 3.5. (46) is a consequence of (45).  $\square$

Furthermore, Lemma 3.5 enables us to reduce the concept of a degree of entailment of to the concept of entailment in degree 1 (full entailment):

**THEOREM 3.7.** *For a graded theory  $T$  and an implication  $A \Rightarrow B$  we have*

$$\|A \Rightarrow B\|_T = \bigvee \{c \in L \mid \|A \Rightarrow c \otimes B\|_T = 1\}.$$

**PROOF.** Using (iii) of Lemma 3.5, we have

$$\begin{aligned} \|A \Rightarrow B\|_T &= \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow B\|_M = \\ &= \bigvee \{c \in L \mid c \leq \|A \Rightarrow B\|_M \text{ for each } M \in \text{Mod}(T)\} = \bigvee \{c \in L \mid \|A \Rightarrow c \otimes B\|_T = 1\}. \end{aligned}$$

$\square$

Therefore, the concept of a degree of entailment by graded theories may be reduced to that of entailment in degree 1 (full entailment) by crisp theories:



COROLLARY 3.8.  $\|A \Rightarrow B\|_T = \bigvee \{c \in L \mid \|A \Rightarrow c \otimes B\|_{\text{cr}(T)} = 1\}$ , with  $\text{cr}(T)$  defined by (44).

### 3.3. Closure Properties of Models of Graded Implications

In the classic setting, models of theories of implications (equivalently, functional dependencies) are closed under intersections. This enables one to test whether  $A \Rightarrow B$  follows from a theory  $T$  by checking whether  $A \Rightarrow B$  is valid in a single model of  $T$ , namely the least model of  $T$  containing  $A$  [Ganter and Wille 1999; Maier 1983]. In this section we establish the corresponding results for a setting with grades.

Recall from [Belohlavek et al. 2005] that a system  $S \subseteq \mathbf{L}^Y$  of  $\mathbf{L}$ -sets in  $Y$  is called an  $\mathbf{L}^*$ -closure system if it is closed under intersections and  $a^*$ -shifts, i.e. satisfies the following conditions:

$$\text{if } A_j \in S \text{ for } j \in J \text{ then } \bigcap_{j \in J} A_j \in S, \quad (47)$$

$$\text{if } a \in L \text{ and } A \in S \text{ then } a^* \rightarrow A \in S. \quad (48)$$

Note that here,  $(\bigcap_{j \in J} A_j)(y) = \bigwedge_{j \in J} A_j(y)$  and  $(a^* \rightarrow A)(y) = a^* \rightarrow A(y)$ . Recall furthermore that an  $\mathbf{L}^*$ -closure operator [Belohlavek et al. 2005] on a set  $Y$  is a mapping  $C : \mathbf{L}^Y \rightarrow \mathbf{L}^Y$  satisfying, for each  $A, A_1, A_2 \in \mathbf{L}^Y$ ,

$$A \subseteq C(A), \quad (49)$$

$$S(A_1, A_2)^* \leq S(C(A_1), C(A_2)), \quad (50)$$

$$C(A) = C(C(A)), \quad (51)$$

where  $S$  is the degree of inclusion defined by (36). If  $L = \{0, 1\}$ ,  $\mathbf{L}^*$ -closure systems and  $\mathbf{L}^*$ -closure operators may be identified with ordinary closure systems and closure operators [Davey and Priestly 2002], since (48) is satisfied for free and (50) asserts monotony of  $C$  with respect to set inclusion in this case. According to [Belohlavek et al. 2005], letting for an  $\mathbf{L}^*$ -closure system  $S$  and an  $\mathbf{L}^*$ -closure operator  $C$ ,

$$C_S(B) = \bigcap_{i \in I} (S(B, A_i)^* \rightarrow A_i) \quad (52)$$

and

$$S_C = \{A \in \mathbf{L}^U \mid A = C(A)\},$$

$C_S$  is an  $\mathbf{L}^*$ -closure operator,  $S_C$  is an  $\mathbf{L}^*$ -closure system, and the mappings  $S \mapsto C_S$  and  $C \mapsto S_C$  are mutually inverse bijections.

THEOREM 3.9. *Mod( $T$ ) is an  $\mathbf{L}^*$ -closure system in  $Y$  for any graded theory  $T$  of implications over  $Y$ .*

PROOF. We need to check (47) and (48). Due to Theorem 3.6, we may safely assume that  $T$  is crisp.

(47): Consider a  $J$ -indexed system  $\{M_j \in \text{Mod}(T) \mid j \in J\}$  of models of  $T$ . We show that  $\bigcap_{j \in J} M_j$  is a model of  $T$ . Thus, we check that, for each  $A \Rightarrow B \in T$ ,  $\|A \Rightarrow B\|_{\bigcap_{j \in J} M_j} = 1$ . Since each  $M_j$  is a model of  $T$ , we have  $\|A \Rightarrow B\|_{M_j} = 1$ , i.e.  $S(A, M_j)^* \leq S(B, M_j)$ , for any  $A \Rightarrow B \in T$ . Now, since  $(\bigwedge_{j \in J} a_j)^* \leq \bigwedge_{j \in J} a_j^*$ , we get

$$S(A, \bigcap_{j \in J} M_j)^* = (\bigwedge_{j \in J} S(A, M_j))^* \leq \bigwedge_{j \in J} S(A, M_j)^* \leq \bigwedge_{j \in J} S(B, M_j) = S(B, \bigcap_{j \in J} M_j),$$

proving  $\|A \Rightarrow B\|_{\bigcap_{j \in J} M_j} = 1$ , and hence  $\bigcap_{j \in J} M_j \in \text{Mod}(T)$ .

(48): Let  $M \in \text{Mod}(T)$  and  $a \in L$ . We need to check that  $a^* \rightarrow M$  belongs to  $\text{Mod}(T)$ . Since  $M$  is a model of  $T$ , for each  $A \Rightarrow B \in T$  we have  $\|A \Rightarrow B\|_M = 1$ , i.e.,  $S(A, M)^* \leq S(B, M)$ . Using Lemma 3.5 (i), (31), (32), and monotony of  $\rightarrow$  in the second argument, we get

$$\begin{aligned} S(A, a^* \rightarrow M)^* &= (a^* \rightarrow S(A, M))^* \leq \\ &\leq a^{**} \rightarrow S(A, M)^* = a^* \rightarrow S(A, M)^* \leq a^* \rightarrow S(B, M) = S(B, a^* \rightarrow M), \end{aligned}$$

establishing  $\|A \Rightarrow B\|_{a^* \rightarrow M} = 1$  for an arbitrary  $A \Rightarrow B \in T$ , whence  $a^* \rightarrow M \in \text{Mod}(T)$ .  $\square$

The following assertion shows the converse claim to Theorem 3.9.

THEOREM 3.10. *Let  $S$  be an  $\mathbf{L}^*$ -closure system in  $Y$ . Then there exists a theory  $T$  of graded attribute implications over  $Y$  such that  $S = \text{Mod}(T)$ .*

PROOF. Put  $T = \{A \Rightarrow C_S(A) \mid A \in \mathbf{L}^Y\}$ . Let  $M \in S$ . Then  $M = C_S(M)$  and due to (50),  $S(A, M)^* \leq S(C_S(A), C_S(M)) = S(C_S(A), M)$ , which gives  $\|A \Rightarrow C_S(A)\|_M = 1$ , i.e.  $M$  is a model of  $T$ . This proves  $S \subseteq \text{Mod}(T)$ . Conversely let  $M \notin S$ , i.e.  $M \neq C_S(M)$ . Then  $M \subset C_S(M)$  by (49), whence  $S(C_S(M), M) \neq 1$ . As a result,

$$\|M \Rightarrow C_S(M)\|_M = S(M, M)^* \rightarrow S(C_S(M), M) = 1^* \rightarrow S(C_S(M), M) = S(C_S(M), M) \neq 1,$$

i.e.  $M \notin \text{Mod}(T)$ , proving  $\text{Mod}(T) \subseteq S$ .  $\square$

Theorem 3.9 and Theorem 3.10 imply that systems of models of graded attribute implications over  $Y$  coincide with  $\mathbf{L}^*$ -closure systems over  $Y$ . In particular, given a theory  $T$  and an arbitrary  $A \in \mathbf{L}^Y$ , one may consider the least model of  $T$  that contains  $A$ . As is well-known from the ordinary case [Ganter and Wille 1999], an ordinary implication  $A \Rightarrow B$  follows from  $T$  if and only if the least model of  $T$  that contains  $A$  includes  $B$ . As we show next, this property generalizes to the setting involving grades in that degree of entailment equals degree of inclusion. In our setting, the least model is  $C_{\text{Mod}(T)}(A)$  where  $C_{\text{Mod}(T)}$  is the  $\mathbf{L}^*$ -closure operator corresponding to  $T$  according to Theorem 3.10. As shown by the following theorem,  $C_{\text{Mod}(T)}(A)$  may be used to determine the degree to which  $A \Rightarrow B$  semantically follows from  $T$ . Namely, the degree equals the degree to which  $A \Rightarrow B$  is valid in  $C_{\text{Mod}(T)}(A)$  as well as the degree of inclusion of  $B$  in  $C_{\text{Mod}(T)}(A)$ .

**THEOREM 3.11.** *For every theory  $T$  and a graded attribute implication  $A \Rightarrow B$ ,*

$$\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)} = S(B, C_{\text{Mod}(T)}(A)).$$

PROOF. Clearly,  $\|A \Rightarrow B\|_T \leq \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)}$  because  $C_{\text{Mod}(T)}(A) \in \text{Mod}(T)$ . Moreover, since  $C_{\text{Mod}(T)}$  satisfies (49),

$$\begin{aligned} \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)} &= S(A, C_{\text{Mod}(T)}(A))^* \rightarrow S(B, C_{\text{Mod}(T)}(A)) \\ &= 1 \rightarrow S(B, C_{\text{Mod}(T)}(A)) = S(B, C_{\text{Mod}(T)}(A)). \end{aligned}$$

Take any  $M \in \text{Mod}(T)$ . Due to (50) and  $M = C_{\text{Mod}(T)}(M)$ ,

$$\begin{aligned} S(B, C_{\text{Mod}(T)}(A)) \otimes S(A, M)^* &\leq S(B, C_{\text{Mod}(T)}(A)) \otimes S(C_{\text{Mod}(T)}(A), C_{\text{Mod}(T)}(M)) \leq \\ &\leq S(B, C_{\text{Mod}(T)}(M)) = S(B, M). \end{aligned}$$

Applying adjointness, we get

$$S(B, C_{\text{Mod}(T)}(A)) \leq S(A, M)^* \rightarrow S(B, M) = \|A \Rightarrow B\|_M,$$

for each  $M \in \text{Mod}(T)$ . Hence,  $S(B, C_{\text{Mod}(T)}(A)) \leq \|A \Rightarrow B\|_T$ .  $\square$

### 3.4. Related structures and alternative formulas for validity

Every table  $\langle X, Y, I \rangle$  with grades induces an important pair of operators. These operators, along with the sets of their fixpoints, were studied in [Belohlavek 1999; Belohlavek 2004; Belohlavek and Vychodil 2012]. In this section, we present the basic connections of these structures to graded attribute implications. In addition, we provide alternative formulas for validity of implications.

Given a table  $\langle X, Y, I \rangle$ , consider the operators  $\uparrow : \mathbf{L}^X \rightarrow \mathbf{L}^Y$  and  $\downarrow : \mathbf{L}^Y \rightarrow \mathbf{L}^X$  given by

$$A^\uparrow(y) = \bigwedge_{x \in X} (A(x)^* \rightarrow I(x, y)) \quad \text{and} \quad B^\downarrow(x) = \bigwedge_{y \in Y} (B(y) \rightarrow I(x, y)). \quad (53)$$

The pair  $\langle \uparrow, \downarrow \rangle$  forms an  $\mathbf{L}^*$ -Galois connection [Belohlavek 1999; Belohlavek 2004; Belohlavek and Vychodil 2012]. Note that the formulas in (53) are not symmetric because we consider only a particular form of these operators, which are directly linked to graded attribute implications. The general formulas involve two hedges, one for  $X$  and one for  $Y$ . Using basic rules of predicate fuzzy logic,  $A^\uparrow(y)$  is the truth degree of “for each  $x \in X$ : if it is very true that  $x$  belongs to  $A$  then  $y$  applies to  $x$ ”. Likewise,  $B^\downarrow(x)$  is the truth degree of “for each  $y \in Y$ : if  $y$  belongs to  $B$  then  $y$  applies to  $x$ ”. The set

$$\mathcal{B}(X^*, Y, I) = \{\langle A, B \rangle \in \mathbf{L}^X \times \mathbf{L}^Y \mid A^\uparrow = B, B^\downarrow = A\}$$

of all fixpoints of  $\langle \uparrow, \downarrow \rangle$  is called the *concept lattice* of  $\langle X, Y, I \rangle$ . Its elements, called *formal concepts* of  $\langle X, Y, I \rangle$ , are naturally interpreted as concepts in the sense of traditional logic [Ganter and Wille 1999]. Namely, for every formal concept  $\langle A, B \rangle \in \mathcal{B}(X^*, Y, I)$ ,  $A$  and  $B$  may be seen as its *extent* and its *intent*, i.e. the collections of objects and attributes, respectively,

which are covered by the concept. Both  $A$  and  $B$  are  $\mathbf{L}$ -sets, i.e. represent graded collections and apply to objects and attributes to degrees, not necessarily 0 and 1 only. The set of all intents, which plays an important role for graded attribute implications, is denoted by  $\text{Int}(X^*, Y, I)$ , i.e.

$$\text{Int}(X^*, Y, I) = \{B \in \mathbf{L}^Y \mid \langle A, B \rangle \in \mathcal{B}(X^*, Y, I) \text{ for some } A\}.$$

Note that

$$\text{for each } B \in \mathbf{L}^Y : B \in \text{Int}(X^*, Y, I) \text{ if and only if } B = B^{\downarrow\uparrow} \quad (54)$$

As the following theorem shows, validity of  $A \Rightarrow B$  in a data table  $\langle X, Y, I \rangle$  may equivalently be expressed as the validity of  $A \Rightarrow B$  in the set of all intents of  $\langle X, Y, I \rangle$  as well as the degree of inclusion of  $B$  in the  $\uparrow\downarrow$ -closure of  $A$ .

**THEOREM 3.12.** *For every  $\langle X, Y, I \rangle$ ,*

$$\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} = S(B, A^{\downarrow\uparrow}). \quad (55)$$

**PROOF.** First, we check  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$ . Observe that  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} \leq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$  iff for each  $M \in \text{Int}(X^*, Y, I)$  we have  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} \leq \|A \Rightarrow B\|_M$ , i.e.

$$\bigwedge_{x \in X} (S(A, I_x)^* \rightarrow S(B, I_x)) \leq S(A, M)^* \rightarrow S(B, M).$$

As  $I_x(y) = I(x, y)$ , we have  $S(A, I_x) = \bigwedge_{y \in Y} (A(y) \rightarrow I_x(y)) = A^\downarrow(x)$ . Therefore, the last inequality is equivalent to

$$\bigwedge_{x \in X} (A^\downarrow(x)^* \rightarrow B^\downarrow(x)) \leq S(A, M)^* \rightarrow S(B, M),$$

i.e. to

$$S(A^{\downarrow*}, B^\downarrow) = \bigwedge_{x \in X} ((A^{\downarrow*}(x) \rightarrow B^\downarrow(x)) \leq S(A, M)^* \rightarrow S(B, M),$$

which is equivalent to

$$S(A, M)^* \otimes S(A^{\downarrow*}, B^\downarrow) \leq S(B, M) \quad (56)$$

due to adjointness of  $\otimes$  and  $\rightarrow$ . Thus, it suffices to prove (56) for each  $M \in \text{Int}(X^*, Y, I)$ . For this purpose, consider the operator  $\uparrow$ , the “unhedged” version of  $\uparrow$  defined by

$$A^\uparrow(y) = \bigwedge_{x \in X} (A(x) \rightarrow I(x, y)).$$

The pair  $\langle \uparrow, \downarrow \rangle$  forms an  $\mathbf{L}$ -Galois connection and hence satisfies  $S(C_1, C_2) \leq S(C_2^\uparrow, C_1^\uparrow)$ ,  $S(D_1, D_2) \leq S(D_2^\downarrow, D_1^\downarrow)$ , and  $D \subseteq D^{\downarrow\uparrow}$ , see [Belohlavek 1999]. Due to (54) and since  $S(C, D) \otimes S(D, E) \leq S(C, E)$ , we obtain

$$\begin{aligned} S(A, M)^* \otimes S(A^{\downarrow*}, B^\downarrow) &\leq S(M^\downarrow, A^\downarrow)^* \otimes S(A^{\downarrow*}, B^\downarrow) \leq \\ &\leq S(M^{\downarrow*}, A^{\downarrow*}) \otimes S(A^{\downarrow*}, B^\downarrow) \leq S(M^{\downarrow*}, B^\downarrow) \leq \\ &\leq S(B^{\downarrow\uparrow}, M^{\downarrow*}) = S(B^{\downarrow\uparrow}, M^{\downarrow\uparrow}) = S(B^{\downarrow\uparrow}, M) \leq S(B, M), \end{aligned}$$

verifying (56) and thus  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} \leq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$ . To check  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} \geq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$  it is sufficient to observe that for each  $x \in X$ ,  $I_x \in \text{Int}(X^*, Y, I)$ . This fact follows from (54) since, as one can easily see,

$$I_x = \{1/x\}^\uparrow = \{1/x\}^{*\uparrow} = \{1/x\}^{*\uparrow\downarrow*} = I_x^{\downarrow*} = I_x^{\downarrow\uparrow}.$$

Second, we check  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = S(B, A^{\downarrow\uparrow})$ . We have

$$\begin{aligned} \|A \Rightarrow B\|_{\langle X, Y, I \rangle} &= \\ &= \bigwedge_{x \in X} (S(A, I_x)^* \rightarrow S(B, I_x)) = \\ &= \bigwedge_{x \in X} (A^\downarrow(x)^* \rightarrow B^\downarrow(x)) = \\ &= \bigwedge_{x \in X} (A^{\downarrow*}(x) \rightarrow \bigwedge_{y \in Y} (B(y) \rightarrow I(x, y))) = \\ &= \bigwedge_{y \in Y} \bigwedge_{x \in X} (A^{\downarrow*}(x) \rightarrow (B(y) \rightarrow I(x, y))) = \\ &= \bigwedge_{y \in Y} \bigwedge_{x \in X} (B(y) \rightarrow (A^{\downarrow*}(x) \rightarrow I(x, y))) = \\ &= \bigwedge_{y \in Y} (B(y) \rightarrow \bigwedge_{x \in X} (A^{\downarrow*}(x) \rightarrow I(x, y))) = \\ &= \bigwedge_{y \in Y} (B(y) \rightarrow A^{\downarrow\uparrow}(y)) = S(B, A^{\downarrow\uparrow}), \end{aligned}$$

proving the claim.  $\square$

We now present several other formulas expressing the degree  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$ . They show that, in a sense, globalization may be regarded as the basic hedge in the definition (37). First, for a hedge  $*$  on  $\mathbf{L}$  put  $\text{fix}(*) = \{a \in L \mid a^* = a\}$  (set of all fixpoints of  $*$ ). Furthermore, for  $\bullet, * : L \rightarrow L$  put  $\bullet \leq *$  iff  $a^\bullet \leq a^*$  for each  $a \in L$  ( $\bullet$  is as strong or stronger than  $*$ ). One can easily see that for hedges  $*$  and  $\bullet$  on a complete residuated lattice  $\mathbf{L}$ ,

$$\bullet \leq * \quad \text{iff} \quad \text{fix}(\bullet) \subseteq \text{fix}(*). \quad (57)$$

Denote by  $\|A \Rightarrow B\|_{\bullet}$  the degree of validity of  $A \Rightarrow B$  in  $\dots$  that involves the hedge  $\bullet$ . Thus,  $\|A \Rightarrow B\|_M^\bullet = S(A, M)^\bullet \rightarrow S(B, M)$  and the like. Omitting the superscript, i.e.  $\|A \Rightarrow B\|_M$  always means  $\|A \Rightarrow B\|_M^*$ . We need the following lemma.

**LEMMA 3.13.** *For  $A, B, M \in \mathbf{L}^Y$ , and hedges  $\bullet$  and  $*$  for which  $\bullet \leq *$  we have*

$$\|A \Rightarrow B\|_M = \bigwedge_{a \in L} (S(a^* \otimes A, M)^\bullet \rightarrow S(a^* \otimes B, M)) = S(S(A, M)^* \otimes B, M).$$

**PROOF.**  $\|A \Rightarrow B\|_M = S(A, M)^* \rightarrow S(B, M) = S(S(A, M)^* \otimes B, M)$  follows directly from Lemma 3.5.

Next, we check both inequalities of  $\|A \Rightarrow B\|_M = \bigwedge_{a \in L} (S(a^* \otimes A, M)^\bullet \rightarrow S(a^* \otimes B, M))$ . “ $\leq$ ” is true iff for each  $a \in L$  we have  $S(a^* \otimes A, M)^\bullet \otimes \|A \Rightarrow B\|_M \leq S(a^* \otimes B, M)$  and since  $S(a^* \otimes B, M) = a^* \rightarrow S(B, M)$ , the latter inequality is equivalent to

$$a^* \otimes S(a^* \otimes A, M)^\bullet \otimes \|A \Rightarrow B\|_M \leq S(B, M)$$

which is true. Indeed,

$$\begin{aligned} a^* \otimes S(a^* \otimes A, M)^\bullet \otimes \|A \Rightarrow B\|_M &\leq a^* \otimes S(a^* \otimes A, M)^* \otimes \|A \Rightarrow B\|_M = \\ &= a^* \otimes (a^* \rightarrow S(A, M))^* \otimes \|A \Rightarrow B\|_M \leq a^* \otimes (a^* \rightarrow S(A, M)^*) \otimes \|A \Rightarrow B\|_M \leq \\ &\leq S(A, M)^* \otimes (S(A, M)^* \rightarrow S(B, M)) \leq S(B, M). \end{aligned}$$

To check “ $\geq$ ”, observe that

$$\begin{aligned} \bigwedge_{a \in L} (S(a^* \otimes A, M)^\bullet \rightarrow S(a^* \otimes B, M)) &\leq (\text{put } a = S(A, M)) \\ &\leq S(S(A, M)^* \otimes A, M)^\bullet \rightarrow S(S(A, M)^* \otimes B, M) = \\ &= 1^\bullet \rightarrow S(S(A, M)^* \otimes B, M) = S(A, M)^* \rightarrow S(B, M) = \|A \Rightarrow B\|_M. \end{aligned}$$

$\square$

**THEOREM 3.14.** *For a data table  $\langle X, Y, I \rangle$  with grades, hedges  $\bullet$  and  $*$  with  $\bullet \leq *$ , and a graded attribute attribute implication  $A \Rightarrow B$ ,*

$$\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \bigwedge_{a \in L} \|a^* \otimes A \Rightarrow a^* \otimes B\|_{\langle X, Y, I \rangle}^\bullet \quad (58)$$

$$\bigwedge_{a \in L} \|a^* \otimes A \Rightarrow a^* \otimes B\|_{\langle X, Y, I \rangle} = \quad (59)$$

$$\bigwedge_{a \in L} \|A \Rightarrow B\|_{\langle X, Y, a^* \rightarrow I \rangle}^\bullet = \quad (60)$$

$$\|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}^\bullet. \quad (61)$$

**PROOF.** (58): Since  $\{1/x\}^\uparrow = I_x$  and

$$\|a^* \otimes A \Rightarrow a^* \otimes B\|_{\langle X, Y, I \rangle}^\bullet = \bigwedge_{x \in X} (S(a^* \otimes A, \{1/x\}^\uparrow)^\bullet \rightarrow S(a^* \otimes B, \{1/x\}^\uparrow)),$$

the fact that  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  equals (58) follows directly from Lemma 3.13 and the definition of  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$ .

(59): The expression is a particular case of (58) for  $\bullet = *$ .

(60): Since

$$\begin{aligned} S(a^* \otimes C, I_x) &= \bigwedge_{y \in Y} ((a^* \otimes C(y)) \rightarrow I(x, y)) = \bigwedge_{y \in Y} (C(y) \rightarrow (a^* \rightarrow I(x, y))) = \\ &= S(C, a^* \rightarrow I_x) = S(C, (a^* \rightarrow I)_x), \end{aligned}$$

where  $a^* \rightarrow I$  is the  $a^*$ -shift of  $I$ , i.e.  $(a^* \rightarrow I)(x, y) = a^* \rightarrow I(x, y)$ , we get

$$\begin{aligned} \bigwedge_{a \in L} \|a^* \otimes A \Rightarrow a^* \otimes B\|_{\langle X, Y, I \rangle}^\bullet &= \bigwedge_{a \in L, x \in X} (S(a^* \otimes A, I_x)^\bullet \rightarrow S(a^* \otimes B, I_x)) = \\ &= \bigwedge_{a \in L, x \in X} (S(A, (a^* \rightarrow I)_x)^\bullet \rightarrow S(B, (a^* \rightarrow I)_x)) = \bigwedge_{a \in L} \|A \Rightarrow B\|_{\langle X, Y, a^* \rightarrow I \rangle}^\bullet, \end{aligned}$$

proving that (60) equals (58).

(61): In view of Theorem 3.12, to prove  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle} \leq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}^\bullet$  it suffices to check  $\|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} \leq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}^\bullet$ , which follows from  $S(A, M)^* \rightarrow S(B, M) \leq S(A, M)^\bullet \rightarrow S(B, M)$ . Conversely, since

$$\begin{aligned} \|A \Rightarrow B\|_{\langle X, Y, I \rangle} &= \bigwedge_{a \in L} \|A \Rightarrow B\|_{\langle X, Y, a^* \rightarrow I \rangle}^\bullet = \bigwedge_{x \in X, a \in L} (S(A, (a \rightarrow I)_x)^\bullet \rightarrow S(B, (a \rightarrow I)_x)) = \\ &= \bigwedge_{x \in X, a \in L} (S(A, \{^a/x\}^\uparrow)^\bullet \rightarrow S(B, \{^a/x\}^\uparrow)), \end{aligned}$$

the inequality  $\|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}^\bullet \leq \|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  is equivalent to

$$\bigwedge_{M \in \text{Int}(X^*, Y, I)} (S(A, M)^\bullet \rightarrow S(B, M)) \leq \bigwedge_{x \in X, a \in L} (S(A, \{^a/x\}^\uparrow)^\bullet \rightarrow S(B, \{^a/x\}^\uparrow)),$$

which follows from the fact that  $\{^a/x\}^\uparrow \in \text{Int}(X^*, Y, I)$  for each  $a \in L$  and  $x \in X$ .  $\square$

*Remark 3.15.* (1) We encounter (58) in Section 4 where we prove completeness of certain Armstrong-like rules for graded attribute implications.

(2) The hedge  $\bullet$  of Theorem 3.14 can range in the sense of (57) arbitrarily from globalization, which is the least hedge, up to  $*$  (boundary condition Theorem 3.14). In particular, with  $\bullet$  being the globalization, Theorem 3.14 says that globalization is in a sense, the basic hedge since the degree of validity of  $A \Rightarrow B$  based on a general hedge  $*$  may be expressed as a degree of validity of  $A \Rightarrow B$  that is based on globalization.

#### 4. LOGIC OF GRADED ATTRIBUTE IMPLICATIONS

In this section, we introduce a system for reasoning with graded attribute implications and prove two versions of completeness for this system. In Section 4.1, we prove the ordinary-style completeness, i.e. we prove that a graded attribute implication  $A \Rightarrow B$  is provable from a set  $T$  of implications iff the degree  $\|A \Rightarrow B\|_T$  to which  $A \Rightarrow B$  semantically follows from  $T$  equals 1. In Section 4.2, we present a graded-style completeness theorem. Namely, we introduce the concept of a degree  $|A \Rightarrow B|_T$  of provability of an implication  $A \Rightarrow B$  from an L-set  $T$  of implications and show that  $|A \Rightarrow B|_T = \|A \Rightarrow B\|_T$ , i.e. the degree of provability coincides with the degree of semantic entailment.

##### 4.1. Armstrong-like rules and ordinary-style completeness

Our axiomatic system consists of the following Armstrong-like *deduction rules* [Armstrong 1974].

- (Ax) infer  $A \cup B \Rightarrow A$ ,
  - (Cut) from  $A \Rightarrow B$  and  $B \cup C \Rightarrow D$  infer  $A \cup C \Rightarrow D$ ,
  - (Mul) from  $A \Rightarrow B$  infer  $c^* \otimes A \Rightarrow c^* \otimes B$ ,
- for each  $A, B, C, D \in \mathbf{L}^Y$ , and  $c \in L$ .

*Remark 4.1.* (1) Rules (Ax) and (Cut) are inspired by the well-known ordinary rules of axiom and cut from which they differ in that  $A, B, C, D$  represent L-sets.

(2) Rule (Mul), the rule of multiplication, is a new rule. Note that  $c^* \otimes A$  is defined by  $(c^* \otimes A)(y) = c^* \otimes A(y)$ . If  $*$  is globalization, (Mul) can be omitted. Indeed, for  $c = 1$ , we have  $c^* = 1$  and (Mul) becomes “from  $A \Rightarrow B$  infer  $A \Rightarrow B$ ”, a trivial rule. For  $c < 1$ , we have  $c^* = 0$  and (Mul) becomes “from  $A \Rightarrow B$  infer  $\emptyset \Rightarrow \emptyset$ ” which can be omitted since  $\emptyset \Rightarrow \emptyset$  can be inferred by (Ax).

Provability is defined as usual: A graded attribute implication  $A \Rightarrow B$  is called *provable* from a set  $T$  of implications using a set  $\mathcal{R}$  of deduction rules, written  $T \vdash_{\mathcal{R}} A \Rightarrow B$ , if there is a sequence (a proof)  $\varphi_1, \dots, \varphi_n$  of implications such that  $\varphi_n$  is  $A \Rightarrow B$  and for each  $\varphi_i$  we either have  $\varphi_i \in T$  or  $\varphi_i$  is inferred from some of the preceding formulas using some deduction rule from  $\mathcal{R}$ . If  $\mathcal{R}$  consists of (Ax)–(Mul), we usually omit  $\mathcal{R}$ , and use, for instance,  $T \vdash A \Rightarrow B$  instead of  $T \vdash_{\mathcal{R}} A \Rightarrow B$ .

A deduction rule “from  $\varphi_1, \dots, \varphi_n$  infer  $\varphi$ ”, with graded attribute implications  $\varphi_1, \dots, \varphi_n, \varphi$ , is derivable from a set  $\mathcal{R}$  of deduction rules if  $\{\varphi_1, \dots, \varphi_n\} \vdash_{\mathcal{R}} \varphi$ . The following lemma presents some derived rules (one easily checks that the arguments from the ordinary case apply [Maier 1983]).

LEMMA 4.2. *The following deduction rules are derivable from (Ax) and (Cut):*

- (Ref) infer  $A \Rightarrow A$ ,
- (Wea) from  $A \Rightarrow B$  infer  $A \cup C \Rightarrow B$ ,

(Add) from  $A \Rightarrow B$  and  $A \Rightarrow C$  infer  $A \Rightarrow B \cup C$ ,

(Pro) from  $A \Rightarrow B \cup C$  infer  $A \Rightarrow B$ ,

(Tra) from  $A \Rightarrow B$  and  $B \Rightarrow C$  infer  $A \Rightarrow C$ ,

for each  $A, B, C, D \in \mathbf{L}^Y$ .

A deduction rule “from  $\varphi_1, \dots, \varphi_n$  infer  $\varphi$ ” is *sound* if  $\text{Mod}(\{\varphi_1, \dots, \varphi_n\}) \subseteq \text{Mod}(\{\varphi\})$ , i.e. every model of all of  $\varphi_1, \dots, \varphi_n$  is a model of  $\varphi$ .

LEMMA 4.3. *Each of the deduction rules (Ax)–(Mul) is sound.*

PROOF. The soundness of (Ax) is trivial as  $\|A \cup B \Rightarrow A\|_M = 1$  holds for any  $M \in \mathbf{L}^Y$ . If  $M \in \text{Mod}(\{A \Rightarrow B, B \cup C \Rightarrow D\})$ , i.e.  $S(A, M)^* \leq S(B, M)$  and  $S(B \cup C, M)^* \leq S(D, M)$ , then using  $S(P \cup Q, R) = S(P, R) \wedge S(Q, R)$  and  $(a \wedge b)^* \leq a^* \wedge b^*$  we get

$$\begin{aligned} S(A \cup C, M)^* &= (S(A, M) \wedge S(C, M))^* = (S(A, M) \wedge S(C, M))^* \leq \\ &\leq (S(A, M)^* \wedge S(C, M)^*)^* \leq (S(B, M) \wedge S(C, M))^* = S(B \cup C, M)^* \leq S(D, M), \end{aligned}$$

proving the soundness of (Cut). Let  $\|A \Rightarrow B\|_M = 1$ . Putting  $X = \{x\}$  and  $I(x, y) = M(y)$  for each  $y \in Y$ , (58) yields  $\|A \Rightarrow B\|_M \leq \|a^* \otimes A \Rightarrow a^* \otimes B\|_M$ , and hence  $\|a^* \otimes A \Rightarrow a^* \otimes B\|_M = 1$ , for every  $a \in L$ , proving the soundness of (Mul).  $\square$

A set  $T$  of graded attribute implications is

– *syntactically closed* if for every  $A \Rightarrow B$ ,  $T \vdash A \Rightarrow B$  implies  $A \Rightarrow B \in T$ ,

– *semantically closed* if for every  $A \Rightarrow B$ ,  $\|A \Rightarrow B\|_T = 1$  implies  $A \Rightarrow B \in T$ .

Clearly,  $T$  is syntactically closed iff  $T = \{A \Rightarrow B \mid T \vdash A \Rightarrow B\}$ ; analogously,  $T$  is semantically closed iff  $T = \{A \Rightarrow B \mid \|A \Rightarrow B\|_T = 1\}$ .

LEMMA 4.4. *Let  $T$  be a set of graded attribute implications. If  $T$  is semantically closed then  $T$  is syntactically closed.*

PROOF. Follows from Lemma 4.3 by standard arguments.  $\square$

LEMMA 4.5. *Let  $T$  be a set of graded attribute implications and let both  $Y$  and  $L$  be finite. If  $T$  is syntactically closed then  $T$  is semantically closed.*

PROOF. Let  $T$  be syntactically closed. We need to show that if  $\|A \Rightarrow B\|_T = 1$  then  $A \Rightarrow B \in T$ . We prove this by verifying that if  $A \Rightarrow B \notin T$  then  $\|A \Rightarrow B\|_T \neq 1$ . Let thus  $A \Rightarrow B \notin T$ . Note that since  $T$  is syntactically closed,  $T$  is closed w.r.t. the rules (Ref)–(Tra) of Lemma 4.2.

To see that  $\|A \Rightarrow B\|_T \neq 1$ , we show that there exists a model of  $T$  that is not a model of  $A \Rightarrow B$ . For this purpose, consider  $M = A^+$  where  $A^+$  is the largest  $L$ -set such that  $A \Rightarrow A^+ \in T$ . Note that  $A^+$  exists. Namely,  $S = \{C \mid A \Rightarrow C \in T\}$  is non-empty since  $A \Rightarrow A \in T$  by (Ref),  $S$  is finite by finiteness of  $Y$  and  $L$ , and for  $A \Rightarrow C_1, \dots, A \Rightarrow C_n \in T$ , we have  $A \Rightarrow \bigcup_{i=1}^n C_i \in T$  by a repeated use of (Add).

We now check that (a)  $A^+$  is a model of  $T$  and that (b)  $A^+$  is not a model of  $A \Rightarrow B$ .

(a): Let  $C \Rightarrow D \in T$ . We need to show  $\|C \Rightarrow D\|_{A^+} = 1$ , i.e.  $S(C, A^+)^* \rightarrow S(D, A^+) = 1$  which is equivalent to  $S(C, A^+)^* \otimes D \subseteq A^+$  due to adjointness of  $\otimes$  and  $\rightarrow$ . Since  $A^+$  is the largest one for which  $A \Rightarrow A^+ \in T$ , in order to verify  $S(C, A^+)^* \otimes D \subseteq A^+$ , it is sufficient to show that  $A \Rightarrow S(C, A^+)^* \otimes D \in T$ . We claim (a1)  $A \Rightarrow A^+ \in T$ , (a2)  $A^+ \Rightarrow S(C, A^+)^* \otimes C \in T$ , and (a3)  $S(C, A^+)^* \otimes C \Rightarrow S(C, A^+)^* \otimes D \in T$ . Indeed,  $A \Rightarrow A^+ \in T$  by definition of  $A^+$ .  $A^+ \Rightarrow S(C, A^+)^* \otimes C \in T$  is an instance of (Ax) because  $S(C, A^+)^* \otimes C \subseteq A^+$ , which follows from

$$\begin{aligned} (S(C, A^+)^* \otimes C)(y) &\leq C(y) \otimes S(C, A^+) = C(y) \otimes \bigwedge_{z \in Y} (C(z) \rightarrow A^+(z)) \leq \\ &\leq C(y) \otimes C(y) \rightarrow A^+(y) \leq A^+(y). \end{aligned}$$

Finally,  $S(C, A^+)^* \otimes C \Rightarrow S(C, A^+)^* \otimes D \in T$  by applying (Mul) to  $C \Rightarrow D \in T$ . Now, (Tra) applied to (a1), (a2), and (a3) yields  $A \Rightarrow S(C, A^+)^* \otimes D \in T$ , proving (a).

(b): We need to show  $\|A \Rightarrow B\|_{A^+} \neq 1$ . Note that

$$\|A \Rightarrow B\|_{A^+} = S(A, A^+)^* \rightarrow S(B, A^+) = 1 \rightarrow S(B, A^+) = S(B, A^+).$$

Therefore, if  $\|A \Rightarrow B\|_{A^+} = 1$ , one has  $1 = S(B, A^+)$ , whence  $B \subseteq A^+$ . Since  $A \Rightarrow A^+ \in T$ , (Pro) would give  $A \Rightarrow B \in T$ , a contradiction to the assumption.  $\square$

The following theorem is the ordinary-style completeness theorem of (Ax)–(Mul).

**THEOREM 4.6.** *Let  $L$  and  $Y$  be finite. For a set  $T$  be of graded attribute implications and a graded attribute implication  $A \Rightarrow B$ ,*

$$T \vdash A \Rightarrow B \quad \text{iff} \quad \|A \Rightarrow B\|_T = 1.$$

**PROOF.** Let  $\text{syn}(T) = \{A \Rightarrow B \mid T \vdash A \Rightarrow B\}$  and  $\text{sem}(T) = \{A \Rightarrow B \mid \|A \Rightarrow B\|_T = 1\}$ . We need to show  $\text{syn}(T) = \text{sem}(T)$ . One easily checks that  $\text{syn}(T)$  and  $\text{sem}(T)$  are the least syntactically and semantically closed sets of graded attribute implications that contain  $T$ , respectively. As  $\text{syn}(T)$  is syntactically closed, it is also semantically closed by Lemma 4.5 which means  $\text{sem}(\text{syn}(T)) = \text{syn}(T)$ . Therefore, since  $T \subseteq \text{syn}(T)$ , we get

$$\text{sem}(T) \subseteq \text{sem}(\text{syn}(T)) = \text{syn}(T).$$

In a similar manner, using Lemma 4.4, we get  $\text{syn}(T) \subseteq \text{sem}(T)$ , showing  $\text{syn}(T) = \text{sem}(T)$ .  $\square$

Note that as is well-known, (Ax) and (Cut) form a syntactico-semantically complete system in the ordinary case (i.e. with fuzzy sets replaced by ordinary sets). The system consisting of (Ax), (Cut), and (Mul) results by adding a new rule, (Mul), to a (Ax) and (Cut). In this perspective, (Mul) is the rule that handles intermediate degrees. Alternatively, one could modify (Cut) and use

(Cut') from  $A \Rightarrow e \otimes B$  and  $B \cup C \Rightarrow D$  infer  $A \cup C \Rightarrow e^* \otimes D$

instead of adding (Mul). Namely:

**LEMMA 4.7.** (Ax), (Cut), and (Mul) are equivalent to (Ax) and (Cut').

**PROOF.** First, we show that (Cut') is derivable from (Ax), (Cut), and (Mul). Let  $\vdash A \Rightarrow e \otimes B$  and  $\vdash B \cup C \Rightarrow D$ . Then  $\vdash e^* \otimes (B \cup C) \Rightarrow e^* \otimes D$  by (Mul),  $\vdash (e \otimes B) \cup C \Rightarrow e^* \otimes D$  by (Wea), and  $\vdash A \cup C \Rightarrow e^* \otimes D$  by (Cut').

Conversely, since (Cut) is an instance of (Cut') for  $e = 1$ , it suffices to show that (Mul) is derivable from (Ax) and (Cut'). Since  $c^* \otimes A \Rightarrow c^* \otimes A$  is an instance of (Ax'), we get  $c^* \otimes A \Rightarrow c^{**} \otimes B$  by (Cut') applied on  $c^* \otimes A \Rightarrow c^* \otimes A$  and  $A \Rightarrow B$ ; (32) gives that  $c^* \otimes A \Rightarrow c^{**} \otimes B$  equals  $c^* \otimes A \Rightarrow c^* \otimes B$  which is the desired formula.  $\square$

In the setting which involves grades, the degree  $S(A, B)$  of inclusion of the L-set  $A$  in the L-set  $B$ , as defined by (36) is an important concept generalizing the classical set inclusion. Another one, generalizing set equality, is the degree  $A \approx B$  of equality of  $A$  and  $B$ , defined by

$$A \approx B = \bigwedge_{y \in Y} (A(y) \leftrightarrow B(y)),$$

where  $a \leftrightarrow b = (a \rightarrow b) \wedge (b \rightarrow a)$  is the biresiduum of  $a$  and  $b$ . Note that  $A \approx B$  is a truth degree of the proposition “for each  $y \in Y$ :  $y$  belongs to  $A$  iff  $y$  belongs to  $B$ ” and that  $\approx$  is an L-equivalence relation [Belohlavek 2002; Gottwald 2001; Hájek 1998]. Therefore,  $A \approx B$  may be conceived as a degree of similarity of  $A$  and  $B$ . Both  $S(A, B)$  and  $A \approx B$  naturally enter derived rules, as illustrated by the following lemma.

**LEMMA 4.8.** *The following deduction rules are derivable from (Ax)–(Mul):*

(S) from  $A \Rightarrow B$  infer  $C \Rightarrow S(A, C)^* \otimes B$ ,

(Sub) from  $A \Rightarrow B$  infer  $C \Rightarrow D \otimes S(A, C)^* \otimes S(D, B)$ ,

(Sim) from  $A \Rightarrow B$  infer  $C \Rightarrow D \otimes (A \approx C)^* \otimes (D \approx B)$ ,

for each  $A, B, C, D \in \mathbf{L}^Y$ .

**PROOF.** One may easily check that (S), (Sub), and (Sim) are all sound rules. The assertion then follows from completeness of (Ax)–(Mul).  $\square$

## 4.2. Graded-style completeness

The ordinary-style completeness does not capture semantic entailment to its full extent in that it only provides a syntactic characterization of entailment to degree 1. In this section, we provide a graded-style completeness theorem which extends to general degrees of entailment. Note that the graded-style completeness goes back to Pavelka’s seminal work [Pavelka 1979] and is further worked out, e.g., in [Gerla 2001; Hájek 1998]. It is based on Goguen’s idea [Goguen 1968–9] of a proof as a sequence of weighted formulas, i.e. pairs  $\langle \varphi, a \rangle$  where  $\varphi$  is a formula and  $a$  a truth degree to which  $\varphi$  has been inferred using deduction rules that operate on both formulas and truth degrees. Graded-style completeness then says that the possibly intermediate degree of entailment of  $\varphi$  equals the degree of provability of  $\varphi$ , i.e. the supremum of  $a$ s that appear in  $\langle \varphi, a \rangle$  at the end of proofs.

In our treatment, this agenda is not employed. Instead, we utilize Corollary 3.8 and ordinary-style completeness and define the notion of provability degree accordingly. Namely, for a fuzzy set  $T$  of graded attribute implications and for  $A \Rightarrow B$  we define the degree  $|A \Rightarrow B|_T \in L$  to which  $A \Rightarrow B$  is provable from  $T$  by

$$|A \Rightarrow B|_T = \bigvee \{c \in L \mid \text{cr}(T) \vdash A \Rightarrow c \otimes B\}, \quad (62)$$

where  $\text{cr}(T)$  is defined by (44). Alternatively,  $|A \Rightarrow B|_T$  may be defined as the largest  $c$  for which  $\text{cr}(T) \vdash A \Rightarrow c \otimes B$ :

**LEMMA 4.9.** *The set  $\{c \in L \mid \text{cr}(T) \vdash A \Rightarrow c \otimes B\}$  in (62) has a largest element.*

**PROOF.** Due to Theorem 4.6, we prove the claim by checking that if  $\|A \Rightarrow c_k \otimes B\|_{\text{cr}(T)} = 1$  for  $k \in K$ , then  $\|A \Rightarrow (\bigvee_{k \in K} c_k) \otimes B\|_{\text{cr}(T)} = 1$ . Observe that  $\|A \Rightarrow (\bigvee_{k \in K} c_k) \otimes B\|_{\text{cr}(T)} = 1$  means that for each model  $M \in \text{Mod}(\text{cr}(T))$ , we have  $\|A \Rightarrow (\bigvee_{k \in K} c_k) \otimes B\|_M = 1$ , i.e.  $S(A, M)^* \leq S((\bigvee_{k \in K} c_k) \otimes B, M) = \bigwedge_{y \in Y} ((B(y) \otimes \bigvee_{k \in K} c_k) \rightarrow M(y))$ , which holds iff for every  $y \in Y$ ,  $S(A, M)^* \leq (B(y) \otimes \bigvee_{k \in K} c_k) \rightarrow M(y)$ . Due to adjointness and the distributivity of  $\otimes$  over  $\bigvee$ , the last inequality is equivalent to  $\bigvee_{k \in K} (c_k \otimes B(y) \otimes S(A, M)^*) \leq M(y)$  which holds if for each  $k \in K$ ,  $c_k \otimes B(y) \otimes S(A, M)^* \leq M(y)$ . This is equivalent to  $S(A, M)^* \leq (c_k \otimes B(y)) \rightarrow M(y)$ . Now, the last inequality holds for every  $y \in Y$  iff  $S(A, M)^* \leq S(c_k \otimes B, M)$  which is true due to the assumptions  $\|A \Rightarrow c_k \otimes B\|_{\text{cr}(T)} = 1$  and  $M \in \text{Mod}(\text{cr}(T))$ .  $\square$

We now have:

**THEOREM 4.10.** *Let  $L$  and  $Y$  be finite. Then for every fuzzy set  $T$  of fuzzy attribute implications and  $A \Rightarrow B$  we have*

$$|A \Rightarrow B|_T = \|A \Rightarrow B\|_T.$$

**PROOF.** Consequence of Corollary 3.8 and Theorem 4.6.  $\square$

## 5. BASES OF GRADED ATTRIBUTE IMPLICATIONS

### 5.1. Completeness, non-redundancy, bases

When exploring graded attribute implications of a table  $\langle X, Y, I \rangle$  with grades, one is interested in small informative sets of implications. A reasonable approach is to require, on one hand, that such set contains information about validity in  $\langle X, Y, I \rangle$  of all implications and, on the other hand, is non-redundant. Such sets are investigated in this section.

**Definition 5.1.** A set  $T$  of graded attribute implications is called *complete* in  $\langle X, Y, I \rangle$  if

$$\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{\langle X, Y, I \rangle} \quad (63)$$

for every implication  $A \Rightarrow B$ .

**Remark 5.2.** (1) That is,  $T$  is complete if the degree of entailment by  $T$  coincides with the degree of validity in  $\langle X, Y, I \rangle$ . In this sense, a complete set contains all information about validity in  $\langle X, Y, I \rangle$  via semantic entailment.

(2) Every  $A \Rightarrow B$  from a complete  $T$  is valid in  $\langle X, Y, I \rangle$  to degree 1. This is a direct consequence of (63) and the fact that  $\|A \Rightarrow B\|_T = 1$  for  $A \Rightarrow B \in T$ .

Completeness of  $T$  may be characterized in terms of models of  $T$  the following way:

**THEOREM 5.3.**  *$T$  is complete in  $\langle X, Y, I \rangle$  iff  $\text{Mod}(T) = \text{Int}(X^*, Y, I)$ .*

**PROOF.** Let  $T$  be complete in  $\langle X, Y, I \rangle$ . Let first  $M \in \text{Mod}(T)$ . Due to completeness of  $T$  and Theorem 3.12,

$$\|M \Rightarrow M^{\downarrow\uparrow}\|_T = \|M \Rightarrow M^{\downarrow\uparrow}\|_{\langle X, Y, I \rangle} = S(M^{\downarrow\uparrow}, M^{\downarrow\uparrow}) = 1.$$

As  $M \in \text{Mod}(T)$ ,  $\|M \Rightarrow M^{\downarrow\uparrow}\|_T = 1$  yields  $\|M \Rightarrow M^{\downarrow\uparrow}\|_M = 1$  from which it follows  $1 = S(M, M)^* \leq S(M^{\downarrow\uparrow}, M)$ , i.e.  $M^{\downarrow\uparrow} \subseteq M$ . Since  $M \subseteq M^{\downarrow\uparrow}$  is always the case, we have  $M = M^{\downarrow\uparrow}$ , hence  $M \in \text{Int}(X^*, Y, I)$  by virtue of (54). We proved  $\text{Mod}(T) \subseteq \text{Int}(X^*, Y, I)$ . Conversely, let  $M \in \text{Int}(X^*, Y, I)$ . Clearly, since  $T$  is complete, Theorem 3.12 implies

$$\|A \Rightarrow B\|_M \geq \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} = \|A \Rightarrow B\|_T$$

for every implication  $A \Rightarrow B$ . In particular, if  $A \Rightarrow B \in T$  then  $\|A \Rightarrow B\|_T = 1$  and the last inequality yields  $\|A \Rightarrow B\|_M = 1$ . This shows  $M \in \text{Mod}(T)$  and thus  $\text{Int}(X^*, Y, I) \subseteq \text{Mod}(T)$ .



Conversely, if  $\text{Mod}(T) = \text{Int}(X^*, Y, I)$  then  $T$  is complete since due to Theorem 3.12,

$$\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{\text{Mod}(T)} = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} = \|A \Rightarrow B\|_{\langle X, Y, I \rangle}.$$

□

**Remark 5.4.** (1) Each of the two inclusions of  $\text{Mod}(T) = \text{Int}(X^*, Y, I)$  has a natural meaning. Namely, as an inspection of the proof of Theorem 5.3 shows, (a)  $\text{Mod}(T) \subseteq \text{Int}(X^*, Y, I)$  if and only if  $\|A \Rightarrow B\|_T \geq \|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  for every  $A \Rightarrow B$ , and (b)  $\text{Mod}(T) \supseteq \text{Int}(X^*, Y, I)$  if and only if  $\|A \Rightarrow B\|_T \leq \|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  for every  $A \Rightarrow B$ .

(2) Let  $\text{Mod}(T) \supseteq \{I_x \mid x \in X\}$  with  $I_x$ s given by (42). Then  $\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{\text{Mod}(T)} \leq \|A \Rightarrow B\|_{\langle X, Y, I \rangle}$ . If, on the other hand,  $\|A \Rightarrow B\|_T \leq \|A \Rightarrow B\|_{\langle X, Y, I \rangle}$ , then according to (1)(b),  $\text{Mod}(T) \supseteq \text{Int}(X^*, Y, I)$ , whence also  $\text{Mod}(T) \supseteq \{I_x \mid x \in X\}$  because  $\text{Int}(X^*, Y, I) \supseteq \{I_x \mid x \in X\}$ . This shows that in (1)(b), the condition  $\text{Mod}(T) \supseteq \text{Int}(X^*, Y, I)$  may be replaced by  $\text{Mod}(T) \supseteq \{I_x \mid x \in X\}$ . Now, since  $\text{Mod}(T) \supseteq \{I_x \mid x \in X\}$  says that every  $A \Rightarrow B \in T$  is valid in  $\langle X, Y, I \rangle$  to degree 1, we conclude that in order to check that  $T$  is complete in  $\langle X, Y, I \rangle$ , it suffices to check that every  $A \Rightarrow B \in T$  be valid in  $\langle X, Y, I \rangle$  to degree 1 and that  $\text{Mod}(T) \subseteq \text{Int}(X^*, Y, I)$ .

**Definition 5.5.** A set  $T$  of graded implications is called a *base* of  $\langle X, Y, I \rangle$  if  $T$  is complete in  $\langle X, Y, I \rangle$  and no proper subset of  $T$  is complete in  $\langle X, Y, I \rangle$ .

Alternatively, one can define the notion of a base the following way. Call a set  $T$  of implications *redundant* if there exists  $A \Rightarrow B \in T$  such that  $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} = 1$ . Otherwise, call  $T$  *non-redundant*.

**LEMMA 5.6.** *The following conditions are equivalent:*

- (i)  $T$  is a non-redundant set of implications.
- (ii) For every  $A \Rightarrow B \in T$ :  $\text{Mod}(T) \subset \text{Mod}(T - \{A \Rightarrow B\})$ .
- (iii) For every  $A \Rightarrow B \in T$  there exists  $C \Rightarrow D$  such that  $\|C \Rightarrow D\|_{T - \{A \Rightarrow B\}} < \|C \Rightarrow D\|_T$ .

**PROOF.** (i)  $\Rightarrow$  (ii): As  $\text{Mod}(T) \subseteq \text{Mod}(T - \{A \Rightarrow B\})$  is always the case, it is sufficient to show that  $\text{Mod}(T) \neq \text{Mod}(T - \{A \Rightarrow B\})$ . Indeed, if  $\text{Mod}(T) = \text{Mod}(T - \{A \Rightarrow B\})$ , then

$$\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} = \|A \Rightarrow B\|_{\text{Mod}(T - \{A \Rightarrow B\})} = \|A \Rightarrow B\|_{\text{Mod}(T)} = \|A \Rightarrow B\|_T = 1,$$

a contradiction to non-redundancy of  $T$ .

(ii)  $\Rightarrow$  (iii): Due to (ii), there exists a model  $M$  of  $T - \{A \Rightarrow B\}$  which is not a model of  $A \Rightarrow B$ . Hence, putting  $C \Rightarrow D = A \Rightarrow B$ , we get  $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} < 1 = \|A \Rightarrow B\|_T$ .

(iii)  $\Rightarrow$  (i): If  $T$  were redundant then for some  $A \Rightarrow B \in T$  we have  $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} = 1$  from which it follows  $\text{Mod}(T) \supseteq \text{Mod}(T - \{A \Rightarrow B\})$ . Since the converse inclusion is obvious, we get  $\text{Mod}(T) = \text{Mod}(T - \{A \Rightarrow B\})$ . As a result,  $\|C \Rightarrow D\|_{T - \{A \Rightarrow B\}} = \|C \Rightarrow D\|_T$  for each  $C \Rightarrow D$ , a contradiction to (iii). □

As the following theorem shows, bases are just complete sets that are non-redundant as sets of implications.

**THEOREM 5.7.**  *$T$  is a base of  $\langle X, Y, I \rangle$  if and only if*

- $T$  is complete in  $\langle X, Y, I \rangle$ , and
- $T$  is non-redundant as a set of implications.

**PROOF.** The assertion follows directly from the fact that non-redundancy of  $T$  is equivalent to condition (iii) of Lemma 5.6 and the fact that if  $T$  is complete in  $\langle X, Y, I \rangle$ , one has  $\|C \Rightarrow D\|_T = \|C \Rightarrow D\|_{\langle X, Y, I \rangle}$ . □

Since one is naturally interested in implications that are fully true in data, the following concept is of interest. A set  $T$  of graded attribute implications is called a *1-complete* in  $\langle X, Y, I \rangle$  if

$$\|A \Rightarrow B\|_T = 1 \text{ iff } \|A \Rightarrow B\|_{\langle X, Y, I \rangle} = 1$$

for every implication  $A \Rightarrow B$ . That is, full consequences of a 1-complete set need to be just the implications fully true in  $\langle X, Y, I \rangle$ . Clearly, completeness in  $\langle X, Y, I \rangle$  implies 1-completeness in  $\langle X, Y, I \rangle$ . Interestingly, we have:

**THEOREM 5.8.**  *$T$  is complete in  $\langle X, Y, I \rangle$  iff  $T$  is 1-complete in  $\langle X, Y, I \rangle$ .*

**PROOF.** We need to show that if  $T$  is a 1-complete, it is complete. Let thus  $T$  be 1-complete in  $\langle X, Y, I \rangle$ . Due to Theorem 3.12,  $T$  is complete iff  $\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$ , i.e. iff  $\|A \Rightarrow B\|_{\text{Mod}(T)} = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$  for every  $A \Rightarrow B$ , which we now verify. First, observe that

$$\|A \Rightarrow S(B, A^{\downarrow\uparrow}) \otimes B\|_{\text{Int}(X^*, Y, I)} = 1. \quad (64)$$

Indeed, the equality holds iff for each  $M \in \text{Int}(X^*, Y, I)$ ,

$$S(A, M)^* \leq S(S(B, A^{\downarrow\uparrow}) \otimes B, M).$$

Since  $S(S(B, A^{\downarrow\uparrow}) \otimes B, M) = S(B, A^{\downarrow\uparrow}) \rightarrow S(B, M)$ , the last inequality is equivalent to

$$S(B, A^{\downarrow\uparrow}) \otimes S(A, M)^* \leq S(B, M). \quad (65)$$

Now, due to the properties of  $\langle \uparrow, \downarrow \rangle$  established in [Belohlavek 1999], and due to  $(\bigwedge_{k \in K} a_k)^* \leq \bigwedge_{k \in K} a_k^*$  and (31), we get

$$\begin{aligned} S(A, M)^* &\leq S(A^{\downarrow}, M^{\downarrow})^* = (\bigwedge_{x \in X} (A^{\downarrow}(x) \rightarrow M^{\downarrow}(x)))^* \leq \bigwedge_{x \in X} (A^{\downarrow}(x)^* \rightarrow M^{\downarrow}(x)^*) = \\ &= S(A^{\downarrow*}, M^{\downarrow*}) \leq S(A^{\downarrow* \uparrow}, M^{\downarrow* \uparrow}) = S(A^{\downarrow\uparrow}, M^{\downarrow\uparrow}). \end{aligned}$$

Therefore, using  $M \in \text{Int}(X^*, Y, I)$  and thus  $M = M^{\downarrow\uparrow}$ , we get

$$S(B, A^{\downarrow\uparrow}) \otimes S(A, M)^* \leq S(B, A^{\downarrow\uparrow}) \otimes S(A^{\downarrow\uparrow}, M^{\downarrow\uparrow}) \leq S(B, M^{\downarrow\uparrow}) = S(B, M),$$

verifying (65) and thus also (64). Next, (64), Theorem 3.12, and the assumption that  $T$  is 1-complete yields

$$1 = \|A \Rightarrow S(B, A^{\downarrow\uparrow}) \otimes B\|_{\text{Int}(X^*, Y, I)} = \|A \Rightarrow S(B, A^{\downarrow\uparrow}) \otimes B\|_{\text{Mod}(T)}.$$

Due to (ii) of Lemma 3.5 and (25),

$$\begin{aligned} \|A \Rightarrow S(B, A^{\downarrow\uparrow}) \otimes B\|_{\text{Mod}(T)} &= \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow S(B, A^{\downarrow\uparrow}) \otimes B\|_M = \\ &= \bigwedge_{M \in \text{Mod}(T)} (S(B, A^{\downarrow\uparrow}) \rightarrow \|A \Rightarrow B\|_M) = S(B, A^{\downarrow\uparrow}) \rightarrow \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow B\|_M = \\ &= S(B, A^{\downarrow\uparrow}) \rightarrow \|A \Rightarrow B\|_{\text{Mod}(T)}. \end{aligned}$$

As a result,

$$S(B, A^{\downarrow\uparrow}) \rightarrow \|A \Rightarrow B\|_{\text{Mod}(T)} = 1,$$

i.e. due to (11),  $S(B, A^{\downarrow\uparrow}) \leq \|A \Rightarrow B\|_{\text{Mod}(T)}$ . Since  $S(B, A^{\downarrow\uparrow}) = \|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)}$  due to Theorem 3.12, we established  $\|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} \leq \|A \Rightarrow B\|_{\text{Mod}(T)}$ . The converse inequality,  $\|A \Rightarrow B\|_{\text{Int}(X^*, Y, I)} \geq \|A \Rightarrow B\|_{\text{Mod}(T)}$ , follows directly from  $\text{Int}(X^*, Y, I) \subseteq \text{Mod}(T)$ , which is a consequence of 1-completeness of  $T$  (the same argument as in the proof of Theorem 5.3 applies).  $\square$

## 5.2. Bases and pseudo-intents

A particular type of bases may be obtained from the following collections of  $L$ -sets of attributes.

**Definition 5.9.** A set  $\mathcal{P} \subseteq L^Y$  is called a *system of pseudo-intents* of  $\langle X, Y, I \rangle$  if for each  $P \in L^Y$ :

$$P \in \mathcal{P} \quad \text{iff} \quad P \neq P^{\downarrow\uparrow} \quad \text{and} \quad \|Q \Rightarrow Q^{\downarrow\uparrow}\|_P = 1 \quad \text{for each } Q \in \mathcal{P} \text{ with } Q \neq P. \quad (66)$$

**Remark 5.10.** (a) Recall that  $\|Q \Rightarrow Q^{\downarrow\uparrow}\|_P = 1$  means that  $S(Q, P)^* \leq S(Q^{\downarrow\uparrow}, P)$ . Hence, if  $*$  is the globalization, then since  $S(Q, P)^* = 1$  if  $Q \subseteq P$  and  $S(Q, P)^* = 0$  if  $Q \not\subseteq P$ , condition (66) simplifies to

$$P \in \mathcal{P} \quad \text{iff} \quad P \neq P^{\downarrow\uparrow} \quad \text{and} \quad Q^{\downarrow\uparrow} \subset P \quad \text{for each } Q \in \mathcal{P} \text{ with } Q \subset P. \quad (67)$$

If  $L$  is, moreover, finite then it is easily seen that there exists a unique system of pseudointents of  $\langle X, Y, I \rangle$ . In general, a system of pseudointents is not unique and may even not exist, as we demonstrate below.

(b) For  $L = \{0, 1\}$ , globalization is the only hedge  $*$  on  $L$ . One easily observes that in this case, (5.10) essentially coincides with the definition of a (unique) system of ordinary pseudo-intents [Ganter 1998; Ganter and Wille 1999; Guigues and Duquenne 1986].

The importance of the notion of a system of pseudointents derives from the following theorem.

**THEOREM 5.11.** *If  $\mathcal{P}$  is a system of pseudo-intents of  $\langle X, Y, I \rangle$  then*

$$T = \{P \Rightarrow P^{\downarrow\uparrow} \mid P \in \mathcal{P}\} \quad (68)$$

*is a non-redundant base of  $\langle X, Y, I \rangle$ .*

**PROOF.** First, we show that  $T$  is complete in  $\langle X, Y, I \rangle$ . Due to Theorem 5.3, it is sufficient to show  $\text{Mod}(T) = \text{Int}(X^*, Y, I)$ . To show  $\text{Mod}(T) \subseteq \text{Int}(X^*, Y, I)$ , assume by contradiction that  $M \in \text{Mod}(T) - \text{Int}(X^*, Y, I)$ . As  $M \notin \text{Int}(X^*, Y, I)$ , we have  $M \neq M^{\downarrow\uparrow}$  by (54), thus in particular  $S(M^{\downarrow\uparrow}, M) \neq 1$ . The assumption  $M \in \text{Mod}(T)$  yields that  $\|Q \Rightarrow Q^{\downarrow\uparrow}\|_M = 1$  for every  $Q \in \mathcal{P}$ , which implies  $M \in \mathcal{P}$  by Definition 5.9. Hence,  $M \Rightarrow M^{\downarrow\uparrow}$  belongs to  $T$  and we have

$$\|M \Rightarrow M^{\downarrow\uparrow}\|_M = S(M, M)^* \rightarrow S(M^{\downarrow\uparrow}, M) = 1^* \rightarrow S(M^{\downarrow\uparrow}, M) = S(M^{\downarrow\uparrow}, M) \neq 1,$$

which contradicts  $M \in \text{Mod}(T)$ . To verify  $\text{Mod}(T) \supseteq \text{Int}(X^*, Y, I)$ , observe that if  $M \in \text{Int}(X^*, Y, I)$  then since  $M = M^{\downarrow\uparrow}$  by (54), we have for every  $P \in \mathcal{L}^Y$ , thus in particular for every  $P \in \mathcal{P}$ ,

$$S(P, M)^* = S(M^{\downarrow}, P^{\downarrow})^* \leq S(M^{\downarrow*}, P^{\downarrow*}) \leq S(P^{\downarrow*}, M^{\downarrow*}) = S(P^{\downarrow\uparrow}, M^{\downarrow\uparrow}) = S(P^{\downarrow\uparrow}, M),$$

whence

$$\|P \Rightarrow P^{\downarrow\uparrow}\|_M \leq S(P, M)^* \rightarrow S(P^{\downarrow\uparrow}, M) = 1,$$

establishing that  $M$  is a model of  $T$ .

Second, we check that  $T$  is non-redundant. If  $T' \subset T$ , there exists  $P \in \mathcal{P}$  such that  $P \Rightarrow P^{\downarrow\uparrow} \notin T'$ . Since  $P \in \mathcal{P}$ , Definition 5.9 yields that  $\|Q \Rightarrow Q^{\downarrow\uparrow}\|_P = 1$  for every  $Q \in \mathcal{P}$  with  $Q \neq P$ , whence  $P \in \text{Mod}(T')$ . Since  $\|P \Rightarrow P^{\downarrow\uparrow}\|_P = S(P^{\downarrow\uparrow}, P) \neq 1$  and since  $T$  is complete in  $\langle X, Y, I \rangle$ , we obtain

$$\begin{aligned} \|P \Rightarrow P^{\downarrow\uparrow}\|_{\langle X, Y, I \rangle} &= \|P \Rightarrow P^{\downarrow\uparrow}\|_T = 1 \neq \|P \Rightarrow P^{\downarrow\uparrow}\|_P \geq \\ &\geq \bigwedge_{M \in \text{Mod}(T')} \|P \Rightarrow P^{\downarrow\uparrow}\|_M = \|P \Rightarrow P^{\downarrow\uparrow}\|_{T'}, \end{aligned}$$

establishing that  $T'$  is not complete in  $\langle X, Y, I \rangle$  and hence  $T$  is non-redundant.  $\square$

As the following example illustrates, there may exist multiple systems of pseudo-intents for a given  $\langle X, Y, I \rangle$  which, moreover, vary in size.

**Example 5.12.** Let  $\mathbf{L}$  with  $L = \{0, 0.5, 1\}$  be a Gödel chain with  $*$  being the identity on  $L$ . Consider  $\langle X, Y, I \rangle$ , where  $X = \{x\}$ ,  $Y = \{y, z\}$ , and  $I(x, y) = I(x, z) = 0$ . The following systems of  $\mathbf{L}$ -sets of attributes are the systems of pseudo-intents of  $\langle X, Y, I \rangle$ :

$$\begin{aligned} \mathcal{P}_1 &= \{\{z\}, \{0.5/y, 0.5/z\}, \{y\}\}, & \mathcal{P}_3 &= \{\{y\}, \{0.5/z\}\}, \\ \mathcal{P}_2 &= \{\{z\}, \{0.5/y\}\}, & \mathcal{P}_4 &= \{\{0.5/y\}, \{0.5/z\}\}. \end{aligned}$$

In general, there exist finite data tables  $\langle X, Y, I \rangle$  for which there does not exist any system of pseudo-intents not even if  $*$  is the globalization. This is illustrated by the following example.

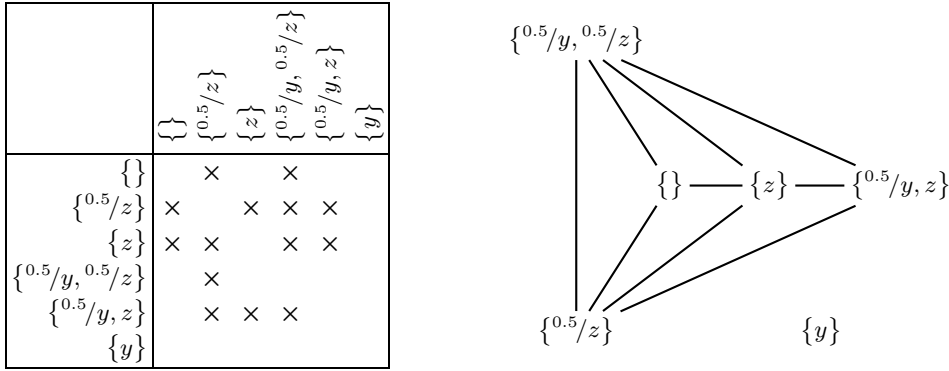
**Example 5.13.** Let  $\mathbf{L}$  be any complete residuated lattice with  $L = [0, 1]$ , let  $*$  be the globalization, and let  $X = \{x\}$ ,  $Y = \{y\}$ , and  $I(x, y) = 0$ . It is easily seen that  $\text{Int}(X^*, Y, I) = \{\{\}, \{y\}\}$ . Assume that there exists a system  $\mathcal{P}$  of pseudo-intents of  $\langle X, Y, I \rangle$ . Due to Theorem 5.11,  $T = \{P \Rightarrow P^{\downarrow\uparrow} \mid P \in \mathcal{P}\}$  is a base and, therefore, Theorem 5.3 implies that  $\text{Mod}(T) = \text{Int}(X^*, Y, I) = \{\{\}, \{y\}\}$ . Thus, for each  $a \in (0, 1)$  there must exist  $\{^c/y\} \in \mathcal{P}$  such that  $\|\{^c/y\} \Rightarrow \{^c/y\}^{\downarrow\uparrow}\|_{\{^a/y\}} \neq 1$ , i.e.

$$(c \rightarrow a)^* = S(\{^c/y\}, \{^a/y\})^* \not\leq S(\{^c/y\}^{\downarrow\uparrow}, \{^a/y\}). \quad (69)$$

Since  $*$  is the globalization, (69) gives  $(c \rightarrow a)^* = 1$ , meaning that  $c \leq a$  and thus  $c \in [0, a]$ . Since  $\{^c/y\}$  is a pseudointent,  $\{^c/y\} \notin \text{Int}(X^*, Y, I) = \{\{\}, \{y\}\}$ , whence  $c \neq 0$  and thus  $c \in (0, a]$ . Now, take any  $b \in (0, c)$ . Repeating the above idea yields a  $d \in (0, b]$  such that  $\{^d/y\} \in \mathcal{P}$  and  $\|\{^d/y\} \Rightarrow \{^d/y\}^{\downarrow\uparrow}\|_{\{^b/y\}} \neq 1$ . Hence, the system of pseudo-intents  $\mathcal{P}$  contains  $\{^c/y\}$  and  $\{^d/y\}$  with  $0 < d < c < 1$ , i.e.  $\{^d/y\} \subset \{^c/y\}$ . However,  $\{^d/y\}^{\downarrow\uparrow} = \{y\} \not\subseteq \{^c/y\}$  which contradicts the assumption that  $\mathcal{P}$  is a system of pseudo-intents.

In the remainder of this section we characterize the systems of pseudo-intents of  $\langle X, Y, I \rangle$  as certain maximal independent sets in graphs associated to  $\langle X, Y, I \rangle$ . For  $\langle X, Y, I \rangle$ , put

$$V = \{P \in \mathbf{L}^Y \mid P \neq P^{\downarrow\uparrow}\}. \quad (70)$$

Fig. 1. Relation  $E$  given by (71) and the induced graph from Example 5.17

Clearly, if  $V$  is empty, then  $\mathcal{P} = \emptyset$  is the only system of pseudo-intents of  $\langle X, Y, I \rangle$ . In this trivial case there is no non-trivial implication valid in  $\langle X, Y, I \rangle$ . For non-empty  $V$  consider the binary relation  $E$  on  $V$  defined by

$$E = \{\langle P, Q \rangle \in V \mid P \neq Q \text{ and } \|Q \Rightarrow Q^{\uparrow\uparrow}\|_P \neq 1\} \quad (71)$$

and the graph  $G = \langle V, E \cup E^{-1} \rangle$ . The following lemma shows a first link between systems of pseudointents and the graph  $G$ .

**LEMMA 5.14.** *If  $\emptyset \neq \mathcal{P}$  is a system of pseudo-intents then  $\mathcal{P}$  is a maximal independent set in  $G$ .*

**PROOF.** Clearly,  $\mathcal{P} \subseteq V$ , since  $P \neq P^{\uparrow\uparrow}$  for every member of a system of pseudo-intents.  $\mathcal{P}$  is independent because otherwise there exist  $P, Q \in \mathcal{P}$  with  $\langle P, Q \rangle \in E$ , i.e.  $\|Q \Rightarrow Q^{\uparrow\uparrow}\|_P \neq 1$ , a contradiction to the definition of a system of pseudointents. **Maximality:** If  $\mathcal{P} \cup \{P\}$  is independent for some  $P \in V - \mathcal{P}$ , then for each  $Q \in \mathcal{P}$  we have  $\langle P, Q \rangle \notin E$ , i.e.  $\|Q \Rightarrow Q^{\uparrow\uparrow}\|_P = 1$ . Definition 5.9 then implies  $P \in \mathcal{P}$ , a contradiction.  $\square$

However, as Example 5.17 shows, there may exist maximal independent sets in  $G$  that are not systems of pseudo-intents. For this reason, define for any  $Q \in V$  and  $\mathcal{P} \subseteq V$  the following subsets of  $V$ :

$$\begin{aligned} \text{Pred}(Q) &= \{P \in V \mid \langle P, Q \rangle \in E\}, \\ \text{Pred}(\mathcal{P}) &= \bigcup_{Q \in \mathcal{P}} \text{Pred}(Q). \end{aligned}$$

The following characterization of systems of pseudo-intents in terms of  $\text{Pred}(\mathcal{P})$  may then be obtained.

**LEMMA 5.15.** *Let  $\mathcal{P} \subseteq V$ .  $\mathcal{P}$  is a system of pseudo-intents iff  $\mathcal{P} = V - \text{Pred}(\mathcal{P})$ .*

**PROOF.** Since every element  $P$  of any system of pseudo-intents satisfies  $P \in V$ , Definition 5.9 implies that being a system of pseudo-intents is equivalent to the following condition:

$$\text{for every } P \in V: P \in \mathcal{P} \text{ iff for each } Q \in \mathcal{P} - \{P\} \text{ we have } \langle P, Q \rangle \notin E.$$

Since  $\langle P, P \rangle \notin E$ , the last condition is equivalent to

$$\text{for every } P \in V: P \in \mathcal{P} \text{ iff } P \notin \text{Pred}(\mathcal{P}),$$

which is clearly equivalent to  $\mathcal{P} = V - \text{Pred}(\mathcal{P})$ .  $\square$

Lemma 5.14 and Lemma 5.15 finally yield the following characterization of systems of pseudo-intents:

**THEOREM 5.16.**  *$\mathcal{P} \neq \emptyset$  is a system of pseudo-intents iff  $\mathcal{P}$  is a maximal independent set in  $G$  such that  $\mathcal{P} = V - \text{Pred}(\mathcal{P})$ .*

Using Theorem 5.16, one may compute systems of pseudo-intents by computing maximal independent sets in  $G$  and checking the additional condition  $\mathcal{P} = V - \text{Pred}(\mathcal{P})$ . Note that this property may be checked when generating the independent sets. The following example illustrates the procedure.

**Example 5.17.** Let  $L$  be a three-element Łukasiewicz chain with  $L = \{0, 0.5, 1\}$ , and  $*$  being the identity on  $L$ . Consider the data table  $\langle X, Y, I \rangle$  where  $X = \{x\}$ ,  $Y = \{y, z\}$ ,  $I(x, y) =$

0.5, and  $I(x, z) = 0$ . The set  $V$  defined by (70) is the following:

$$V = \{\{\}, \{^{0.5}/z\}, \{z\}, \{^{0.5}/y, ^{0.5}/z\}, \{^{0.5}/y, z\}, \{y\}\}.$$

The corresponding binary relation  $E$  defined by (71) is depicted in Fig. 1 (left); graph  $G = \langle V, E \cup E^{-1} \rangle$  is depicted in Fig. 1 (right).  $G$  contains four maximal independent sets:

$$\begin{aligned} \mathcal{P}_1 &= \{\{\}, \{^{0.5}/y, z\}, \{y\}\}, & \mathcal{P}_3 &= \{\{z\}, \{y\}\}, \\ \mathcal{P}_2 &= \{\{^{0.5}/z\}, \{y\}\}, & \mathcal{P}_4 &= \{\{^{0.5}/y, ^{0.5}/z\}, \{y\}\}. \end{aligned}$$

Observe that  $\mathcal{P}_1$  and  $\mathcal{P}_3$  do not satisfy  $\mathcal{P}_i = V - \text{Pred}(\mathcal{P}_i)$  ( $i \in \{1, 3\}$ ) because  $\{^{0.5}/y, ^{0.5}/z\} \notin \text{Pred}(\mathcal{P}_1)$ ,  $i = 1, 3$ , and  $\{\} \notin \text{Pred}(\mathcal{P}_3)$ . Hence, due to Theorem 5.15,  $\mathcal{P}_1$  and  $\mathcal{P}_3$  are not systems of pseudo-intents. On the other hand,  $\mathcal{P}_i = V - \text{Pred}(\mathcal{P}_i)$  for  $i = 2, 4$ , i.e.  $\mathcal{P}_2$  and  $\mathcal{P}_4$  are systems of pseudo-intents of  $\langle X, Y, I \rangle$ . The corresponding non-redundant bases  $T_2$  and  $T_4$  of  $\langle X, Y, I \rangle$  given by Theorem 5.11 are the following:

$$\begin{aligned} T_2 &= \{\{^{0.5}/z\} \Rightarrow \{y, ^{0.5}/z\}, \{y\} \Rightarrow \{y, ^{0.5}/z\}\}, \\ T_4 &= \{\{^{0.5}/y, ^{0.5}/z\} \Rightarrow \{y, ^{0.5}/z\}, \{y\} \Rightarrow \{y, ^{0.5}/z\}\}. \end{aligned}$$

Further algorithmic aspects of this procedure are discussed in Section 6.

### 5.3. Pseudo-intents and bases corresponding to globalization

It has been pointed out in Remark 5.10 that if the hedge  $*$  involved in the definition of the validity of attribute implications is the globalization and if  $L$  and  $Y$  are finite, there exists a unique system of pseudointents for a given  $\langle X, Y, I \rangle$ . In this section, we show that in this case, the corresponding bases are minimal in terms of the number of implications contained in the base. For the subsequent proofs, we need the following technical observation which applies to general systems of pseudointents using any hedge.

**LEMMA 5.18.** *Let  $\mathcal{P}$  be a system of pseudointents of  $\langle X, Y, I \rangle$  and let  $P, Q \in \mathcal{P} \cup \text{Int}(X^*, Y, I)$  satisfy*

$$S(P, Q)^* \leq S(P^{\downarrow\uparrow}, P \cap Q), \quad (72)$$

$$S(Q, P)^* \leq S(Q^{\downarrow\uparrow}, P \cap Q). \quad (73)$$

*Then  $P \cap Q \in \text{Int}(X^*, Y, I)$ .*

**PROOF.** Put  $T' = T - \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$ , where  $T$  is a set of fuzzy attribute implications defined by (68). Definition 5.9 and the fact that  $\|C \Rightarrow C^{\downarrow\uparrow}\|_D = 1$  for every  $C \in \mathbf{L}^Y$  and  $D \in \text{Int}(X^*, Y, I)$  imply  $P, Q \in \text{Mod}(T')$ . Hence, for each  $A \Rightarrow B \in T'$  we have  $S(A, P)^* \leq S(B, P)$  and  $S(A, Q)^* \leq S(B, Q)$ . Consequently,

$$\begin{aligned} S(A, P \cap Q)^* &= (S(A, P) \wedge S(A, Q))^* \leq S(A, P)^* \wedge S(A, Q)^* \leq \\ &\leq S(B, P) \wedge S(B, Q) = S(B, P \cap Q), \end{aligned}$$

which yields that  $P \cap Q$  is a model of  $T'$ . Due to Theorem 5.3, it is now sufficient to verify that  $P \cap Q$  is a model of  $\{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$ . By virtue of (72) and (73), we have

$$S(P, P \cap Q)^* = S(P, Q)^* \leq S(P^{\downarrow\uparrow}, P \cap Q)$$

and

$$S(Q, P \cap Q)^* = S(Q, P)^* \leq S(Q^{\downarrow\uparrow}, P \cap Q),$$

i.e.  $\|P \Rightarrow P^{\downarrow\uparrow}\|_{P \cap Q} = 1$  and  $\|Q \Rightarrow Q^{\downarrow\uparrow}\|_{P \cap Q} = 1$ , finishing the proof.  $\square$

**Remark 5.19.** If  $P$  and  $Q$  are intents or pseudo-intents satisfying  $S(P, Q)^* = S(Q, P)^* = 0$  then (72) and (73) are met and due to Lemma 5.18,  $P \cap Q$  is an intent. Hence, if  $*$  is the globalization and  $P$  and  $Q$  are intents or pseudo-intents with  $P \not\subseteq Q$  and  $Q \not\subseteq P$ , then  $P \cap Q$  is an intent.

**THEOREM 5.20.** *Let  $L$  be a finite residuated lattice with  $*$  being the globalization, let  $Y$  be finite. Let  $\mathcal{P}$  be the sytem of pseudo-intents of  $\langle X, Y, I \rangle$  and  $T$  be the corresponding base given by (68). Then for any base  $T'$  of  $\langle X, Y, I \rangle$  we have  $|T| \leq |T'|$ .*

**PROOF.** We first show that for each  $P \in \mathcal{P}$ ,  $T'$  contains an implication  $A \Rightarrow B$  such that  $A \subseteq P$  and  $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$ . We then show that two distinct  $P, Q \in \mathcal{P}$  cannot share the same implication satisfying this property which proves that  $|T| = |\mathcal{P}| \leq |T'|$ .

Take any  $P \in \mathcal{P}$ . By definition,  $P \neq P^{\downarrow\uparrow}$  and thus  $P \notin \text{Int}(X^*, Y, I)$ . Since  $T'$  is a base, Theorem 5.3 yields that  $T'$  contains  $A \Rightarrow B$  such that  $\|A \Rightarrow B\|_P \neq 1$ . Since  $*$  is the globalization,

we get  $A \subseteq P$  and  $B \not\subseteq P$ . As every implication in  $T'$  is valid in  $\langle X, Y, I \rangle$  to degree 1, Theorem 3.12 yields  $S(B, A^{\downarrow\uparrow}) = 1$ , i.e.  $B \subseteq A^{\downarrow\uparrow}$ . Thus, from  $B \subseteq A^{\downarrow\uparrow}$  and  $B \not\subseteq P$  it follows that  $A^{\downarrow\uparrow} \not\subseteq P$ . Now,  $A \subseteq P$  and  $A^{\downarrow\uparrow} \not\subseteq P$  yield  $A \subseteq A^{\downarrow\uparrow} \cap P \subset A^{\downarrow\uparrow}$ . Since  $A^{\downarrow\uparrow}$  is the least intent containing  $A$ , it follows that  $A^{\downarrow\uparrow} \cap P$  is not an intent. Next, we claim that  $P \subseteq A^{\downarrow\uparrow}$ . By contradiction, if  $P \not\subseteq A^{\downarrow\uparrow}$  then since  $A^{\downarrow\uparrow} \not\subseteq P$ , Lemma 5.18 would give  $A^{\downarrow\uparrow} \cap P \in \text{Int}(X^*, Y, I)$ , a contradiction to the above observation that  $A^{\downarrow\uparrow} \cap P \notin \text{Int}(X^*, Y, I)$ . Now,  $A \subseteq P$  yields  $A^{\downarrow\uparrow} \subseteq P^{\downarrow\uparrow}$  while  $P \subseteq A^{\downarrow\uparrow}$  yields  $P^{\downarrow\uparrow} \subseteq A^{\downarrow\uparrow\downarrow\uparrow} = A^{\downarrow\uparrow}$ , showing  $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$ .

Now, consider  $P, Q \in \mathcal{P}$  such that  $P \neq Q$  and assume that  $A \Rightarrow B \in T'$  satisfies  $A \subseteq P$ ,  $A \subseteq Q$ , and  $P^{\downarrow\uparrow} = A^{\downarrow\uparrow} = Q^{\downarrow\uparrow}$ . If  $P \subset Q$ , then  $P^{\downarrow\uparrow} \subset Q$  and thus  $A^{\downarrow\uparrow} = P^{\downarrow\uparrow} \subset Q \subset Q^{\downarrow\uparrow} = A^{\downarrow\uparrow}$ , a contradiction. Dually for  $Q \subset P$ . Thus, assume that  $P \not\subseteq Q$  and  $Q \not\subseteq P$ . Using Lemma 5.18, we get  $P \cap Q \in \text{Int}(X^*, Y, I)$  and using the assumption that  $A \subseteq P$  and  $A \subseteq Q$ , it follows that  $A \subseteq P \cap Q$ , i.e.,  $A^{\downarrow\uparrow} \subseteq (P \cap Q)^{\downarrow\uparrow} = P \cap Q \subset P^{\downarrow\uparrow}$ , a contradiction.  $\square$

## 6. ALGORITHMS

In this section, we present algorithms for computing bases. We start by an algorithm which simplifies the graph-theoretic procedure based on Theorem 5.16 from Section 5.2. In case of globalization, we can show that the maximal independent set which determines the (uniquely given) system of pseudo-intents can be directly computed without the need to go over all maximal independent sets of the graph. A simplified algorithm which follows is based on the following observation.

**THEOREM 6.1.** *Let  $\mathbf{L}$  be a finite linearly ordered residuated lattice with  $*$  being the globalization and let  $\sqsubset$  be a strict total order on  $\mathbf{L}^Y$  which extends  $\subset$ . Furthermore, assume that  $\mathcal{P}$ ,  $V$ , and  $E$  are given by (66), (70), and (71), respectively. Let for  $P \in \mathcal{P}$  denote  $\mathcal{Q} = \{Q \in \mathcal{P} \mid Q \sqsubset P\}$ . Then  $P$  is the least element of  $(V - \mathcal{Q}) - \text{Pred}(\mathcal{Q})$  with respect to  $\sqsubset$ .*

**PROOF.** First, we prove that  $P \in (V - \mathcal{Q}) - \text{Pred}(\mathcal{Q})$ . Obviously,  $P \in V - \mathcal{Q}$  and thus it suffices to check that  $P \notin \text{Pred}(\mathcal{Q})$  which means showing  $P \not\subseteq \text{Pred}(Q)$  for all  $Q \in \mathcal{Q}$ . Since  $*$  is globalization,  $P \not\subseteq \text{Pred}(Q)$  and  $P \neq Q$  mean that  $Q^{\downarrow\uparrow} \subset P$  whenever  $Q \subset P$  which is indeed true because  $P \in \mathcal{P}$ , cf. (67). Second, we prove that  $P$  is the least element of  $(V - \mathcal{Q}) - \text{Pred}(\mathcal{Q})$ . Assume by contradiction that  $P' \sqsubset P$  for some  $P' \in (V - \mathcal{Q}) - \text{Pred}(\mathcal{Q})$ . Since  $P' \in V - \mathcal{Q}$ , we get  $P' \not\subseteq \mathcal{Q}$ . On the other hand, from  $P' \notin \text{Pred}(\mathcal{Q})$  it follows that  $Q^{\downarrow\uparrow} \subseteq P'$  whenever  $Q \in \mathcal{P}$  and  $Q \subset P'$  which by (67) gives  $P' \in \mathcal{P}$  and thus  $P' \sqsubset P$  gives  $P' \in \mathcal{Q}$ , a contradiction.  $\square$

---

### ALGORITHM 1: Computing the systems of pseudo-intents (case of globalization)

---

**Data:**  $\langle X, Y, I \rangle$  (input data),  $S$  (list of  $\mathbf{L}$ -sets  $\{P \mid P \neq P^{\downarrow\uparrow}\}$  sorted by  $\sqsubset$ )

**Result:**  $\mathcal{P}$  (subset of  $\mathbf{L}^Y$ )

$\mathcal{P} := \emptyset;$

**while** IsNotEmpty( $S$ ) **do**

$\mathcal{P} := \mathcal{P} \cup \{\text{First}(S)\};$

$\mathcal{R} := \text{NewList}();$

$B := \text{First}(S);$

$S := \text{Rest}(S);$

**while** IsNotEmpty( $S$ ) **do**

**if**  $B \subset \text{First}(S)$  **and**  $B^{\downarrow\uparrow} \not\subseteq \text{First}(S)$  **then**

$\text{Put}(\mathcal{R}, \text{First}(S));$

**end**

$S := \text{Rest}(S);$

**end**

$S := \mathcal{R};$

**end**

**return**  $\mathcal{P};$

---

Directly from Theorem 6.1, we derive a procedure for computing pseudo-intents which utilizes the observation that in order to compute  $P \in \mathcal{P}$ , it suffices to find all pseudo-intents which are strictly smaller than  $P$  according to a strict total  $\sqsubset$  order extending  $\subset$ . The procedure is formalized as Algorithm 1. The algorithm involves the following operations with linked lists:  $\text{First}(S)$  (the first element of list  $S$ ),  $\text{Rest}(S)$  (the rest of the list  $S$  except for the first element),  $\text{Put}(S, B)$  (destructive modification of the list  $S$  by putting the element  $B$  to its end),  $\text{IsNotEmpty}(S)$  (condition true if list  $S$  is not empty),  $\text{NewList}()$  (a constructor for a

new empty list). The algorithm takes  $\langle X, Y, I \rangle$  as the input and a list  $S$  which consists of all elements of  $V$  given by (70) which are put in the list in the order according to  $\sqsubset$ .

**THEOREM 6.2.** *If  $L$  is a finite linear residuated lattice and  $*$  is globalization, then Algorithm 1 is correct: For  $\langle X, Y, I \rangle$  and  $S$  which is a list of elements  $P \in L^Y$  satisfying  $P \neq P^{\downarrow\uparrow}$  which are sorted according to a total strict order  $\sqsubset$  extending  $\subset$ , the algorithm terminates after finitely many steps and it returns  $P$  satisfying (66).*

**PROOF.** It is easily seen that the algorithm always terminates because we consecutively remove elements from the list  $S$  and it eventually becomes empty. By induction on the number of loops of the outer while-loop, it suffices to check that whenever the algorithm reaches the beginning of the loop body,  $P$  contains all the elements from (67) which are smaller than  $\text{First}(S)$  according to  $\sqsubset$  provided that  $S$  is nonempty, and  $S$  consists of all elements from  $(V - P) - \text{Pred}(P)$  and that equality  $P \cap \text{Pred}(P) = \emptyset$  is satisfied. The base case is clear. In the induction step, if  $P$  and  $S$  have these properties, from Theorem 6.1 it follows that  $\text{First}(S)$  can be added to  $P$  and the inner while-loop of the algorithm computes new  $S$  which contains all elements of  $(V - Q) - \text{Pred}(Q)$  for  $Q = P \cup \{\text{First}(S)\}$ . Moreover,  $Q \cap \text{Pred}(Q) = \emptyset$  because  $P \in Q \cap \text{Pred}(Q)$  would violate the fact that all elements from  $Q$  are a subset of the elements from (67). So, for the updated  $P$  and  $S$ , the condition holds. Therefore, at the end of the computation,  $S$  is empty, meaning that  $(V - P) - \text{Pred}(P) = \emptyset$ , i.e.,  $V - P \subseteq \text{Pred}(P)$ . Since  $P \cap \text{Pred}(P) = \emptyset$ , we get  $V - P = \text{Pred}(P)$ . Now, apply Theorem 5.15.  $\square$

Algorithm 1 is limited only to globalization  $*$  and does not produce systems of pseudo-intents for general hedges. Although it is more efficient than the naive application of Theorem 5.15 which involves looking for all maximal independent sets, it still uses a large search space which is in general exponential in terms of the size of  $Y$  and  $L$ .

An alternative approach to computing minimal bases using globalization and complete sets using general hedges utilizes the idea of computing fixed points of particular closure operators associated to  $\langle X, Y, I \rangle$ . In particular, for any set  $T$  of graded attribute implications and any  $L$ -set  $C \in L^Y$ , we consider an non-decreasing sequence of  $L$ -sets  $C_1, C_2, \dots$  such that  $C_1 = C$  and

$$C_{i+1} = C_i \cup \bigcup \{B \mid A \Rightarrow B \in T \text{ and } A \subset C_i\}, \quad (74)$$

for any natural number  $i$  and put

$$[C]_T = \bigcup_{n=1}^{\infty} C_n. \quad (75)$$

If  $L$  is finite and linearly ordered and  $Y$  is finite, we get by the Tarski fixed point theorem that  $[\cdot\cdot\cdot]_T$  defined by (75) is a closure operator. In addition, since both  $L$  and  $Y$  are finite,  $[C]_T = C_n$  for some natural  $n$ . Furthermore, we obtain the following theorem:

**THEOREM 6.3.** *Let  $L$  be finite and linearly ordered,  $Y$  be finite,  $\mathcal{P}$  be a system satisfying (67), and let  $T$  be given by (68). Then  $T$  is complete in  $\langle X, Y, I \rangle$  and*

$$\text{fix}([\cdot\cdot\cdot]_T) = \mathcal{P} \cup \text{Int}(X^*, Y, I),$$

i.e.  $C = [C]_T$  iff  $C \in \mathcal{P} \cup \text{Int}(X^*, Y, I)$ .

**PROOF.** The fact that  $T$  is complete in  $\langle X, Y, I \rangle$  can be shown analogously as in the case of Theorem 5.11. We therefore omit the proof but notice here that our  $\mathcal{P}$ , uniquely given by (67) (even if we consider a general hedge), need not satisfy (66). Now, we prove that the set of all fixed points of  $[\cdot\cdot\cdot]_T$  coincides with  $\mathcal{P} \cup \text{Int}(X^*, Y, I)$ .

Let  $P \in \mathcal{P}$  and take  $Q \Rightarrow Q^{\downarrow\uparrow} \in T$  such that  $Q \subset P$ . Directly from (67), we get  $Q^{\downarrow\uparrow} \subseteq P$  and thus  $[P]_T \subseteq P$ , i.e.,  $P$  is a fixed point of  $[\cdot\cdot\cdot]_T$ . Take  $B \in \text{Int}(X^*, Y, I)$  and  $Q \Rightarrow Q^{\downarrow\uparrow} \in T$  such that  $Q \subset B$ . By monotony of  $^{\downarrow\uparrow}$ , we get  $Q^{\downarrow\uparrow} \subseteq B^{\downarrow\uparrow} = B$ , i.e.,  $[B]_T \subseteq B$ . Conversely, let  $C = [C]_T$  such that  $C \neq C^{\downarrow\uparrow}$ . It suffices to check that  $C \in \mathcal{P}$ . Since  $C$  is a fixed point of  $[\cdot\cdot\cdot]_T$ , we get  $Q^{\downarrow\uparrow} \subseteq C$  for any  $Q \Rightarrow Q^{\downarrow\uparrow} \in T$  such that  $Q \subset C$ . Using (67) and (68), the latter gives  $C \in \mathcal{P}$ .  $\square$

Theorem 6.3 can be used to compute both the sets of intents of  $\langle X, Y, I \rangle$  and the set  $\mathcal{P}$  given by (67) for which the set  $T$  given by (68) is complete in  $T$ . In case of the globalization,  $T$  is a minimal base due to Theorem 5.20. A procedure based on this observation is presented in Algorithm 2. In order to simplify notation, attribute sets used in the algorithm are subsets of integers. In the algorithm, we use the following notation: for  $a \in L$  such that  $a < 1$  we denote by  $a^+$  the least element of  $(a, 1]$ . Such  $a^+$  always exists since we assume that  $L$  is a finite and linearly ordered.

**ALGORITHM 2:** Determining intents and complete sets**Data:**  $\langle X, Y, I \rangle$  where  $Y = \{1, \dots, n\}$  (input data)**Result:**  $\mathcal{I}$  and  $\mathcal{P}$  (subsets of  $\mathbf{L}^Y$ ) $\mathcal{P} := \emptyset;$  $\mathcal{I} := \emptyset;$ **if**  $\emptyset = \emptyset^{\uparrow\uparrow}$  **then**|  $\mathcal{I} := \{\emptyset\};$ **else**|  $\mathcal{P} := \{\emptyset\};$ **end** $B := \emptyset;$ **while**  $B \neq Y$  **do**|  $T := \{B \Rightarrow B^{\uparrow\uparrow} \mid B \in \mathcal{P}\};$ |  $C := B;$ | **for**  $y := 1$  **to**  $n$  **do**| | **if**  $C(y) < 1$  **then**| | |  $B := [\{C(y)^{\uparrow}/y, C(y+1)/y+1, \dots, C(n)/n\}]_T;$ | | | **if**  $B(z) = C(z)$  **for all**  $z = 1, \dots, y-1$  **then**

| | | | break for loop;

| | **end**| **end**| **end**| **if**  $B = B^{\uparrow\uparrow}$  **then**| |  $\mathcal{I} := \mathcal{I} \cup \{B\};$ | **else**| |  $\mathcal{P} := \mathcal{P} \cup \{B\};$ | **end****end****return**  $\mathcal{I}, \mathcal{P};$ 

**THEOREM 6.4.** *If  $\mathbf{L}$  is a finite linear residuated lattice, then Algorithm 2 is correct: For  $\langle X, Y, I \rangle$ , the algorithm terminates after finitely many steps and returns  $\mathcal{I}$  and  $\mathcal{P}$  such that  $\mathcal{I} = \text{Int}(X^*, Y, I)$  and  $\mathcal{P}$  satisfies (67).*

**PROOF.** Suppose that  $Y = \{1, \dots, n\}$ . The algorithm maintains two sets:  $\mathcal{I}$  contains only  $\mathbf{L}$ -sets  $B$  such that  $B = B^{\uparrow\uparrow}$  and  $\mathcal{P}$  contains only  $\mathbf{L}$ -sets  $B$  such that  $B \neq B^{\uparrow\uparrow}$ . The main loop of the algorithm goes through all fixed points of  $[\cdot \cdot \cdot]_T$  in the lexicographic order  $\sqsubset$  where  $B_1 \sqsubset B_2$  iff there is  $y \in Y$  such that  $B_1(y) < B_2(y)$  and  $B_1(z) = B_2(z)$  for all  $z > y$  (recall that for simplicity we have denoted attributes as integers). Indeed, the for-loop finds a lexical successor of  $B$  with respect to such  $\sqsubset$  which is a fixed point of  $[\cdot \cdot \cdot]_T$ . Note that  $T$  already contains all necessary implications to compute such fixed point because all elements of (67) which are strictly smaller than the current  $B$  are already in  $\mathcal{P}$ . Hence, at the end of the computation,  $\mathcal{P}$  and  $\mathcal{I}$  consists of all the fixed points of  $[\cdot \cdot \cdot]_T$ . The rest follows from Theorem 6.3.  $\square$

**Remark 6.5.** (1): If  $*$  is a general hedge then Algorithm 2 produces  $\mathcal{P}$  such that the corresponding theory  $T$  given by (68) is complete but may be redundant. In order to get a non-redundant one, i.e. a base, we may consecutively remove from  $T$  graded implications which follow from other graded implications from the theory, i.e., we may repeatedly apply Lemma 5.6(iii). Namely,  $T$  is non-redundant if there is no  $A \Rightarrow B$  such that  $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} = 1$ . According to Theorem 3.11, the equality can be checked by showing  $B \subseteq C_{\text{Mod}(T - \{A \Rightarrow B\})}(A)$ , i.e., by showing whether  $B$  is contained in the least fixed point of  $C_{\text{Mod}(T - \{A \Rightarrow B\})}$  containing  $A$ .

(2): If  $\mathbf{L}$  and  $Y$  are finite, the fixed points of  $C_{\text{Mod}(T)}$  which play a role in the previous remark can be efficiently computed. Namely, for  $M \in \mathbf{L}^Y$  we may put  $M_1 = M$  and

$$M_{i+1} = M \cup \bigcup \{B \otimes S(A, M)^* \mid A \Rightarrow B \in T\} \quad (76)$$

for any natural number  $i$ . It is easy to see that  $C_{\text{Mod}(T)}(M) = \bigcup_{n=1}^{\infty} M_n$ . Note that since  $L$  and  $Y$  are finite,  $\bigcup_{n=1}^{\infty} M_n$  is equal to  $M_n$  for some  $n$ .

(3): The complexity of computing bases derives from the fact that even for  $L = \{0, 1\}$ , there may exist an exponential number of pseudo-intents in terms of  $|X|$  and  $|Y|$  (number of objects and attributes) [Kuznetsov and Obiedkov 2008]. Hence, since the size of a smallest base equals the number of pseudo-intents, a smallest base may have an exponential size in terms



of  $|X|$  and  $|Y|$  in the worst case. On the other hand, the time delay [Johnson et al. 1988], which is for the above reason an appropriate concept in our case, of Algorithm 2 is  $\leq O(|L|)$ -times the time delay of the basic algorithm for computing ordinary pseudo-intents [Ganter 1984], which is also described in [Ganter and Wille 1999].

## 7. REDUCING GRADED ATTRIBUTE IMPLICATIONS TO ORDINARY ONES VIA THRESHOLDING

As mentioned above, ordinary attribute implications are a particular case of graded implications in which 0 and 1 are the only degrees involved. In this section we look at whether and to what extent the notions regarding graded attribute implications and their bases may be reduced to those regarding ordinary implications. In particular, we show that every data table  $\langle X, Y, I \rangle$  with graded attributes may be transformed via a natural thresholding to a table  $\langle X^\times, Y^\times, I^\times \rangle$  with binary attributes in such a way that validity of graded implications in  $\langle X, Y, I \rangle$  corresponds to validity of the ordinary implications in  $\langle X^\times, Y^\times, I^\times \rangle$ . A natural question arises of whether bases of  $\langle X, Y, I \rangle$  may be obtained from the bases of  $\langle X^\times, Y^\times, I^\times \rangle$ , since the latter ones may be computed by existing algorithms [Ganter and Wille 1999]. As we show, the answer to this question is negative. Namely, while complete sets of ordinary implications in  $\langle X^\times, Y^\times, I^\times \rangle$  yield complete sets of graded implications in  $\langle X, Y, I \rangle$  via the transformation, it may happen that non-redundant sets of ordinary implications transform to redundant sets of graded implications.

The transformation via thresholding is based on the following idea. Given a graded attribute  $y$ , one may consider for every truth degree  $b \in L$  the corresponding bivalent attribute  $\langle y, b \rangle$  as follows:  $\langle y, b \rangle$  applies to the object  $x$  if and only if  $y$  applies to  $x$  at least to degree  $b$ . This idea is, in fact, a particular case of a more general one which underlies the following definition.

Given a table  $\langle X, Y, I \rangle$  with graded attributes, denote by  $\langle X^\times, Y^\times, I^\times \rangle$  the table with graded attributes defined by:

$$\begin{aligned} X^\times &= X \times *(L), \text{ where } *(L) = \{a^* \mid a \in L\}, \\ Y^\times &= Y \times L, \\ \langle \langle x, a \rangle, \langle y, b \rangle \rangle &\in I^\times \text{ iff } a \otimes b \leq I(x, y). \end{aligned}$$

One may easily check using the properties of hedges that  $*(L) = \{a \in L \mid a^* = a\}$ , i.e.  $*(L)$  is the set of all fixpoints of  $*$ . If  $*$  is globalization,  $*(L) = \{0, 1\}$ , the new objects of the form  $\langle x, 1 \rangle$  may be identified with the original objects  $x \in X$  while those of the form  $\langle x, 0 \rangle$  may be dropped because they are redundant (every new attribute  $\langle y, b \rangle$  applies to them). In this case, the ordinary relation  $I^\times$  coincides with the one which corresponds to the simple thresholding as described above because then,  $\langle y, b \rangle$  applies to  $x$ , i.e. to  $\langle x, 1 \rangle$ , iff  $b = 1 \otimes b \leq I(x, y)$ , i.e.  $y$  applies to  $x$  at least to degree  $b$ .

To transform graded attribute implications to ordinary ones and vice versa, we utilize the following mappings between L-sets and ordinary sets. For an L-set  $B \in \mathbf{L}^Y$  we define the ordinary subset  $\lfloor B \rfloor$  of  $Y \times L$  by

$$\lfloor B \rfloor = \{\langle y, a \rangle \in Y \times L \mid a \leq B(y)\}.$$

For a subset  $D \subseteq Y \times L$  we define the L-set  $\lceil D \rceil$  in  $Y$  by

$$\lceil D \rceil(y) = \bigvee \{a \mid \langle y, a \rangle \in D\}.$$

With these correspondences, one may look at the relationship between the validity of graded implications in  $\langle X, Y, I \rangle$  on one hand and the validity of ordinary implications in  $\langle X^\times, Y^\times, I^\times \rangle$  on the other hand. Namely, for a given graded implication  $A \Rightarrow B$  over  $Y$  (i.e.  $A, B \in \mathbf{L}^Y$ ), one may consider the corresponding ordinary implication  $\lfloor A \rfloor \Rightarrow \lfloor B \rfloor$  over  $Y \times L$  (i.e.  $\lfloor A \rfloor, \lfloor B \rfloor \subseteq Y \times L$ ), and conversely, for an ordinary implication  $C \Rightarrow D$  over  $Y \times L$ , one may consider the corresponding graded implication  $\lceil C \rceil \Rightarrow \lceil D \rceil$  over  $Y$ . The relationship in question is described by the following theorem which says that the transformations described above preserve validity of implications (for brevity, we write  $\|A \Rightarrow B\|_I$  instead of  $\|A \Rightarrow B\|_{\langle X, Y, I \rangle}$  and the same for  $\|A \Rightarrow B\|_{I^\times}$ ).

**THEOREM 7.1.** *For a data table  $\langle X, Y, I \rangle$  with graded attributes, the corresponding  $\langle X^\times, Y^\times, I^\times \rangle$ , and arbitrary  $A \in \mathbf{L}^Y$ ,  $B \in \mathbf{L}^Y$  and  $C, D \subseteq Y \times L$ , we have*

$$\|A \Rightarrow B\|_I = 1 \text{ if and only if } \|\lfloor A \rfloor \Rightarrow \lfloor B \rfloor\|_{I^\times} = 1; \quad (77)$$

$$\|C \Rightarrow D\|_{I^\times} = 1 \text{ if and only if } \|\lceil C \rceil \Rightarrow \lceil D \rceil\|_I = 1. \quad (78)$$

Before we turn to the proof of Theorem 7.1, we present some auxiliary results. Denote by  $\wedge : 2^{X \times *(L)} \rightarrow 2^{Y \times L}$  and  $\vee : 2^{Y \times *(L)} \rightarrow 2^{X \times *(L)}$  the Galois connections induced by  $I^\times$  [Ore 1944], i.e.

$$\begin{aligned} C^\wedge &= \{\langle y, b \rangle \in Y \times L \mid \text{for each } \langle x, a \rangle \in C : \langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^\times\}, \text{ and} \\ D^\vee &= \{\langle x, a \rangle \in X \times *(L) \mid \text{for each } \langle y, b \rangle \in D : \langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^\times\}, \end{aligned}$$

for every  $C \subseteq X \times *(L)$  and  $D \subseteq Y \times L$ . Furthermore, let us extend  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  for any  $A \in \mathbf{L}^X$  and  $C \subseteq X \times *(L)$  by putting  $\lfloor A \rfloor = \{\langle x, a \rangle \in X \times *(L) \mid a \leq A(x)\}$  and  $\lceil C \rceil(y) = \bigvee \{a \mid \langle x, a \rangle \in C\}$ . As  $(\bigvee_k a_k^*)^* = \bigvee_k a_k^*$  (due to the isotony and idempotency of  $*$ ),  $*(L)$  is closed under suprema and, hence,  $\lceil C \rceil(x) \in *(L)$  for every  $x \in X$ . The following lemma describes the relationship between  $\langle \uparrow, \downarrow \rangle$  and  $\langle \wedge, \vee \rangle$ , and some further properties.

**LEMMA 7.2.** *For every  $A \in \mathbf{L}^X$ ,  $B \in \mathbf{L}^Y$ ,  $C \subseteq X \times *(L)$ , and  $D \subseteq Y \times L$ ,*

$$A^\uparrow = \lceil \lfloor A^* \rfloor \rceil, \quad B^\downarrow = \lfloor \lceil B \rceil \rfloor, \quad C^\wedge = \lfloor \lceil C \rceil^\uparrow \rfloor, \quad \text{and} \quad D^\vee = \lceil \lfloor D \rfloor^\downarrow \rceil; \quad (79)$$

$$C^\wedge = \lfloor \lceil C \rceil \rfloor^\wedge \quad \text{and} \quad D^\vee = \lceil \lfloor D \rfloor \rceil^\vee; \quad (80)$$

$$\lfloor A \rfloor^{\wedge\vee} = \lfloor A^{\uparrow\downarrow} \rfloor, \quad \lfloor B \rfloor^{\vee\wedge} = \lfloor B^{\downarrow\uparrow} \rfloor, \quad \lceil C \rceil^{\uparrow\downarrow} = \lceil C^{\wedge\vee} \rceil, \quad \text{and} \quad \lceil D \rceil^{\downarrow\uparrow} = \lceil D^{\vee\wedge} \rceil. \quad (81)$$

**PROOF.** (79) We have

$$\begin{aligned} \lceil \lfloor A^* \rfloor \rceil(y) &= \bigvee \{b \mid \langle y, b \rangle \in \lfloor A^* \rfloor\} = \bigvee \{b \mid \text{for each } \langle x, a \rangle \in \lfloor A^* \rfloor : \langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^\times\} = \\ &= \bigvee \{b \mid \text{for each } \langle x, a \rangle \in \lfloor A^* \rfloor : a \otimes b \leq I(x, y)\} = \\ &= \bigvee \{b \mid \text{for each } x \in X, a \in *(L) : a \leq A^*(x) \text{ implies } a \otimes b \leq I(x, y)\} = \\ &= \bigvee \{b \mid \text{for each } x \in X : A^*(x) \otimes b \leq I(x, y)\} = \bigvee \{b \mid b \leq \bigwedge_{x \in X} (A^*(x) \rightarrow I(x, y))\} = \\ &= \bigwedge_{x \in X} (A^*(x) \rightarrow I(x, y)) = A^\uparrow(y), \end{aligned}$$

proving  $A^\uparrow = \lceil \lfloor A^* \rfloor \rceil$ .  $B^\downarrow = \lfloor \lceil B \rceil \rfloor$  is proven analogously.

To verify  $C^\wedge = \lfloor \lceil C \rceil^\uparrow \rfloor$ , we reason as follows:

$$\begin{aligned} \langle y, b \rangle \in \lfloor \lceil C \rceil^\uparrow \rfloor &\text{ iff } b \leq \lceil C \rceil^\uparrow(y) = \bigwedge_{x \in X} (\lceil C \rceil^*(x) \rightarrow I(x, y)) \\ &\text{ iff for each } x \in X : b \leq \lceil C \rceil^*(x) \rightarrow I(x, y) = (\bigvee_{\langle x, a \rangle \in C} a)^* \rightarrow I(x, y) = (\bigvee_{\langle x, a \rangle \in C} a) \rightarrow I(x, y) \\ &\text{ iff for each } x \in X : \bigvee_{\langle x, a \rangle \in C} (a \otimes b) = (\bigvee_{\langle x, a \rangle \in C} a) \otimes b \leq I(x, y) \\ &\text{ iff for each } \langle x, a \rangle \in C : a \otimes b \leq I(x, y) \\ &\text{ iff for each } \langle x, a \rangle \in C : \langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^\times \text{ iff } \langle y, b \rangle \in C^\wedge. \end{aligned}$$

$D^\vee = \lceil \lfloor D \rfloor^\downarrow \rceil$  is proven analogously.

(80): Due to (79) and since  $\lceil \lfloor A \rfloor \rceil = A$  for every  $A \in \mathbf{L}^X$ ,  $C^\wedge = \lfloor \lceil C \rceil^\uparrow \rfloor = \lfloor \lceil \lfloor \lceil C \rceil \rfloor \rceil^\uparrow \rfloor = \lfloor \lceil C \rceil \rfloor^\wedge$ .  $D^\vee = \lceil \lfloor D \rfloor \rceil^\vee$  is proven analogously.

(81): By virtue of (79) and since  $\lceil \lfloor M \rfloor \rceil = M$ , we have  $\lfloor A \rfloor^{\wedge\vee} = \lfloor \lceil \lfloor A \rfloor \rceil^\uparrow \rfloor^\vee = \lfloor \lceil \lfloor \lceil A \rceil \rfloor \rceil^\uparrow \rfloor^\downarrow \rceil = \lfloor A^{\uparrow\downarrow} \rfloor$ . The second equality is proven analogously.

Due to (79) and (80),  $\lceil C \rceil^{\uparrow\downarrow} = \lceil \lfloor \lceil C \rceil \rfloor \rceil^\uparrow \rfloor^\downarrow \rceil = \lceil C^{\wedge\vee} \rceil$ . The last equality is proven dually.  $\square$

**PROOF OF THEOREM 7.1.** (77): Due to Theorem 3.12,  $\|A \Rightarrow B\|_I = 1$  is equivalent to  $S(B, A^{\uparrow\downarrow})$ , i.e. to  $B \subseteq A^{\uparrow\downarrow}$ , and  $\|\lfloor A \rfloor \Rightarrow \lfloor B \rfloor\|_{I^\times} = 1$  is equivalent to  $\lfloor B \rfloor \subseteq \lfloor A \rfloor^{\vee\wedge}$ . Now, due to (81),  $\lfloor A \rfloor^{\vee\wedge} = \lfloor A^{\uparrow\downarrow} \rfloor$ . Hence, we need to check that  $B \subseteq A^{\uparrow\downarrow}$  if and only if  $\lfloor B \rfloor \subseteq \lfloor A^{\uparrow\downarrow} \rfloor$  which is clearly the case since for every  $M, N \in \mathbf{L}^Y$ ,  $M \subseteq N$  is equivalent to  $\lfloor M \rfloor \subseteq \lfloor N \rfloor$ .

(78): We prove the claim by establishing that (a)  $D \subseteq C^{\vee\wedge}$  is equivalent to (b)  $\lfloor D \rfloor \subseteq \lceil C \rceil^{\uparrow\downarrow}$ . Namely, on account of Theorem 3.12, (a) is equivalent to  $\|C \Rightarrow D\|_{I^\times} = 1$  and (b), i.e.  $S(\lfloor D \rfloor, \lceil C \rceil^{\uparrow\downarrow}) = 1$ , is equivalent to  $\|\lceil C \rceil \Rightarrow \lceil D \rceil\|_I = 1$ . Since (a) clearly implies (b), i.e.  $S(\lceil C \rceil, \lceil D \rceil) = 1$ , is equivalent to  $\|\lceil C \rceil \Rightarrow \lceil D \rceil\|_I = 1$ , we see that (a) implies (b). Assume (b). Then clearly,  $\lfloor \lfloor D \rfloor \rfloor \subseteq \lfloor \lceil C \rceil^{\uparrow\downarrow} \rfloor$ . As  $\lfloor \lceil C \rceil^{\uparrow\downarrow} \rfloor = \lfloor \lceil C \rceil \rfloor^{\vee\wedge} = C^{\vee\wedge}$  on account of (81) and (80), we get  $\lfloor \lfloor D \rfloor \rfloor \subseteq C^{\vee\wedge}$ . As  $D \subseteq \lfloor \lfloor D \rfloor \rfloor$ , we have  $D \subseteq C^{\vee\wedge}$ , proving that (b) implies (a).  $\square$

**Remark 7.3.** In addition to (78) of Theorem 7.1, we also have

$$\|C \Rightarrow D\|_{I^\times} = 1 \quad \text{if and only if} \quad \|\lfloor \lceil C \rceil \rfloor \Rightarrow \lfloor \lceil D \rceil \rfloor\|_{I^\times} = 1.$$

Namely, the two conditions involved are equivalent to (a)  $D \subseteq C^{\vee\wedge}$  and (b)  $\lfloor \lfloor D \rfloor \rfloor \subseteq \lfloor \lceil C \rceil \rfloor^{\vee\wedge} = C^{\vee\wedge}$ , respectively, on account of Theorem 3.12 and (80). Since  $D \subseteq \lfloor \lfloor D \rfloor \rfloor$ , (b) and (80) clearly imply (a). On the other hand, (a) implies  $\lfloor \lfloor D \rfloor \rfloor \subseteq \lfloor \lceil C^{\vee\wedge} \rceil \rfloor$ . Since  $C^{\vee\wedge} = \lceil \lceil C \rceil \rfloor^\uparrow$ , we have  $\lfloor \lceil C^{\vee\wedge} \rceil \rfloor = \lfloor \lceil \lceil C \rceil \rfloor^\uparrow \rfloor = \lfloor \lceil C \rceil \rfloor^\vee = C^{\vee\wedge}$ . As a result, (a) implies (b).

In view of the above results, a natural question is whether one can obtain complete sets and bases of a given table  $\langle X, Y, I \rangle$  with graded attributes from complete sets and bases of the corresponding  $\langle X^\times, Y^\times, I^\times \rangle$ . This question is the subject of the next theorem and the following remark.

**THEOREM 7.4.** *If  $T^\times$  is complete in  $\langle X^\times, Y^\times, I^\times \rangle$  then*

$$\lceil T^\times \rceil = \{ \lceil C \rceil \Rightarrow \lceil D \rceil \mid C \Rightarrow D \in T^\times \} \quad (82)$$

*is complete in  $\langle X, Y, I \rangle$ .*

**PROOF.** Let  $T^\times$  be complete in  $\langle X \times *(L), Y \times L, I^\times \rangle$ . Due to Theorem 5.3, it is sufficient to show that  $\text{Mod}(\lceil T^\times \rceil) = \text{Int}(X^*, Y, I)$ . We prove this fact by showing that the following claims are equivalent for any  $M \in \mathbf{L}^Y$ :

- (a)  $M \in \text{Mod}(\lceil T^\times \rceil)$ ,
- (b) for each  $a \in L$ :  $\lfloor a^* \rightarrow M \rfloor \in \text{Mod}(T^\times)$ ,
- (c)  $M \in \text{Int}(X^*, Y, I)$ .

“(a)  $\Leftrightarrow$  (b)”: Clearly, it suffices to show that for every  $C \Rightarrow D \in T^\times$ ,  $M$  is a model of  $\lceil C \rceil \Rightarrow \lceil D \rceil$  iff for each  $a \in L$ ,  $\lfloor a^* \rightarrow M \rfloor$  is a model of  $C \Rightarrow D$ , i.e. that  $S(\lceil C \rceil, M)^* \leq S(\lceil D \rceil, M)$  iff for each  $a \in L$ ,  $C \subseteq \lfloor a^* \rightarrow M \rfloor$  implies  $D \subseteq \lfloor a^* \rightarrow M \rfloor$ .

Observe first that  $C \subseteq \lfloor a^* \rightarrow M \rfloor$  is equivalent to  $a^* \leq S(\lceil C \rceil, M)$ : Namely,  $C \subseteq \lfloor a^* \rightarrow M \rfloor$  means that for every  $y \in Y$ , if  $\langle y, b \rangle \in C$  then  $\langle y, b \rangle \in \lfloor a^* \rightarrow M \rfloor$ , i.e.  $b \leq (a^* \rightarrow M)(y) = a^* \rightarrow M(y)$ . Therefore,  $C \subseteq \lfloor a^* \rightarrow M \rfloor$  means that for every  $y \in Y$ ,  $\bigvee_{\langle y, b \rangle \in C} b \leq a^* \rightarrow M(y)$  which holds iff for every  $y \in Y$ ,

$$a^* \leq (\bigvee_{\langle y, b \rangle \in C} b) \rightarrow M(y) = \lceil C \rceil(y) \rightarrow M(y)$$

which is equivalent to  $a^* \leq \bigwedge_{y \in Y} (\lceil C \rceil(y) \rightarrow M(y)) = S(\lceil C \rceil, M)$ .

Since the same holds for  $D$ , to prove that (a) is equivalent to (b), it is sufficient to check that  $S(\lceil C \rceil, M)^* \leq S(\lceil D \rceil, M)$  iff for every  $a \in L$ ,  $a^* \leq S(\lceil C \rceil, M)$  implies  $a^* \leq S(\lceil D \rceil, M)$ , which is easy to see. Indeed, if  $S(\lceil C \rceil, M)^* \leq S(\lceil D \rceil, M)$  and  $a^* \leq S(\lceil C \rceil, M)$ , then

$$a^* = a^{**} \leq S(\lceil C \rceil, M)^* \leq S(\lceil D \rceil, M).$$

Conversely, putting  $a = S(\lceil C \rceil, M)$  the assumption, i.e.  $a^* \leq S(\lceil C \rceil, M)$  implies  $a^* \leq S(\lceil D \rceil, M)$ , readily yields  $S(\lceil C \rceil, M)^* \leq S(\lceil D \rceil, M)$ .

“(b)  $\Leftrightarrow$  (c)”: Theorem 5.3 (actually, its instance for  $L = \{0, 1\}$ ) implies that  $\lfloor a^* \rightarrow M \rfloor \in \text{Mod}(T^\times)$  iff  $\lfloor a^* \rightarrow M \rfloor \in \text{Int}(\langle X \times *(L), Y \times L, I^\times \rangle)$ .

Next, observe that  $\lfloor N \rfloor \in \text{Int}(\langle X \times *(L), Y \times L, I^\times \rangle)$  is equivalent to  $N \in \text{Int}(X^*, Y, I)$ . (54) implies that to check this, it suffices to check that  $\lfloor N \rfloor = \lfloor N \rfloor^{\vee \wedge}$  is equivalent to  $N = N^{\downarrow \uparrow}$ . Using Lemma 7.2, we reason as follows. If  $\lfloor N \rfloor = \lfloor N \rfloor^{\vee \wedge}$ , then

$$N^{\downarrow \uparrow} = \lceil \lceil \lfloor N \rfloor^{\vee \wedge} \rceil \rceil = \lceil \lfloor N \rfloor^{\vee \wedge} \rceil = \lceil \lfloor N \rfloor \rceil = N.$$

Here, we used that  $\lceil \lceil P \rceil \rceil = P$  for any  $P \in \mathbf{L}^Y$  (obvious) and  $\lceil \lceil \lfloor N \rfloor^{\vee \wedge} \rceil \rceil = \lfloor N \rfloor^{\vee \wedge}$  which holds because  $\lceil \lceil \lfloor N \rfloor^{\vee \wedge} \rceil \rceil = \lceil \lceil \lceil \lfloor N \rfloor^{\vee \wedge} \rceil^{\downarrow} \rceil \rceil = \lceil \lceil \lfloor N \rfloor^{\downarrow} \rceil \rceil = \lfloor N \rfloor$ . Conversely, if  $N = N^{\downarrow \uparrow}$  then

$$\lfloor N \rfloor^{\vee \wedge} = \lceil \lceil \lceil \lfloor N \rfloor^{\downarrow} \rceil^{\uparrow} \rceil \rceil = \lceil N^{\downarrow \uparrow} \rceil = \lfloor N \rfloor.$$

Applying this observation to  $N = a^* \rightarrow M$ , we see that (b) is equivalent to that fact that for each  $a \in L$ :  $a^* \rightarrow M \in \text{Int}(X^*, Y, I)$ . The proof is complete by observing that the fact that for each  $a \in L$  we have  $a^* \rightarrow M \in \text{Int}(X^*, Y, I)$  is equivalent to  $M \in \text{Int}(X^*, Y, I)$ . Indeed, since for  $a = 1$  we have  $a \rightarrow M = M$ , it is sufficient to observe that if  $M \in \text{Int}(X^*, Y, I)$  then  $a^* \rightarrow M \in \text{Int}(X^*, Y, I)$  for each  $a \in L$ . This follows from the fact that  $\text{Int}(X^*, Y, I)$  is an  $\mathbf{L}^*$ -closure system [Belohlavek et al. 2005] (cf. Section 3.3).  $\square$

The set  $\lceil T^\times \rceil$  obtained from a given  $T^\times$  according to Theorem 7.4 need not be a base of  $\langle X, Y, I \rangle$  even if  $T^\times$  is a base of  $\langle X^\times, Y^\times, I^\times \rangle$ . Namely, as the next example shows,  $\lceil T^\times \rceil$  may be redundant.

**Example 7.5.** Let  $\mathbf{L}$  be the three-element Łukasiewicz chain with  $L = \{0, 0.5, 1\}$ , let  $*$  be the globalization, and let  $X = \{x\}$ ,  $Y = \{y, z\}$ ,  $I(x, y) = I(x, z) = 0$ . One may verify that (abbreviating  $\langle y, a \rangle$  by  $y_a$ )

$$\mathcal{P}^\times = \{ \{y_0, y_{0.5}, z_0\}, \{y_0, y_1, z_0\}, \{y_0, z_0, z_{0.5}\}, \{y_0, z_0, z_1\}, \{\} \}$$

is the system of pseudo-intents of  $\langle X^\times, Y^\times, I^\times \rangle$ . Therefore,

$$\begin{aligned} T^\times = & \{ \{y_0, y_{0.5}, z_0\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\}, \\ & \{y_0, y_1, z_0\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\}, \\ & \{y_0, z_0, z_{0.5}\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\}, \\ & \{y_0, z_0, z_1\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\}, \\ & \{\} \Rightarrow \{y_0, z_0\} \} \end{aligned}$$

is a base of  $\langle X^\times, Y^\times, I^\times \rangle$ . Clearly,

$$[T^\times] = \{ \{^{0.5}/y\} \Rightarrow \{y, z\}, \{y\} \Rightarrow \{y, z\}, \{^{0.5}/z\} \Rightarrow \{y, z\}, \{z\} \Rightarrow \{y, z\}, \{\} \Rightarrow \{\} \}.$$

According to Theorem 7.4,  $[T^\times]$  is complete in  $\langle X, Y, I \rangle$ . Now,  $[T^\times]$  is redundant. First,  $[T^\times]$  contains a trivial implication  $\{\} \Rightarrow \{\}$  which holds true in each  $M \in \mathbf{L}^Y$ . Furthermore,  $[T^\times] - \{\{\} \Rightarrow \{\}\}$  is still redundant, because implications  $\{y\} \Rightarrow \{y, z\}$  and  $\{z\} \Rightarrow \{y, z\}$  semantically follow from

$$S = \{ \{^{0.5}/y\} \Rightarrow \{y, z\}, \{^{0.5}/z\} \Rightarrow \{y, z\} \},$$

i.e.  $\|\{y\} \Rightarrow \{y, z\}\|_S = 1$  and  $\|\{z\} \Rightarrow \{y, z\}\|_S = 1$ .

Example 7.5 also shows that the converse claim to that of Theorem 7.4 does not hold. That is, it is not true that if a set  $T$  of graded implications is complete in  $\langle X, Y, I \rangle$  then  $[T] = \{ \lfloor A \rfloor \Rightarrow \lfloor B \rfloor \mid A \Rightarrow B \in T \}$  is complete in  $\langle X^\times, Y^\times, I^\times \rangle$ . Namely, if this were true then for the set  $S$  from Example 7.5, which is complete in  $\langle X, Y, I \rangle$ , the set

$$\lfloor S \rfloor = \{ \{y_0, y_{0.5}, z_0\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\}, \{y_0, z_0, z_{0.5}\} \Rightarrow \{y_0, y_{0.5}, y_1, z_0, z_{0.5}, z_1\} \}$$

would be complete in  $\langle X^\times, Y^\times, I^\times \rangle$  which it is not, because  $\lfloor S \rfloor$  is a proper subset of a base of  $\langle X^\times, Y^\times, I^\times \rangle$ , namely of  $T^\times$ .

The two observations, namely that  $[T^\times]$  may be redundant even when  $T^\times$  is not, and that  $[T]$  need not be complete even when  $T$  is, have the following explanation. The dependencies reflecting the algebraic structure  $\mathbf{L}$  of the set of grades are implicitly taken into account in the definition of entailment of graded implications over  $Y$ , i.e. in the semantics using  $\mathbf{L}$  as the structure of truth degrees, and need not be present in  $T$ . Their counterparts, however, are “not known” to the definition of (bivalent) semantic entailment of ordinary implications over  $Y \times L$ , and need thus be explicitly present in  $T^\times$ .

## 8. RELATIONSHIP TO FUNCTIONAL DEPENDENCIES OVER DOMAINS WITH SIMILARITIES

In this section, we point out a connection between graded attribute implications and certain extensions of Codd’s relational model of data. Recall that in the ordinary case, which corresponds to  $L = \{0, 1\}$  in our setting, the following connection was presented in [Fagin 1977]. Ordinary attribute implications have two basic interpretations, namely, as propositional logic formulas and as functional dependencies. An attribute implication, say  $\{y_1, y_2, y_3\} \Rightarrow \{z_1, z_2\}$ , may be conceived as a logic formula  $y_1 \& y_2 \& y_3 \supset z_1 \& z_2$  in which  $y_i$ s and  $z_j$ s are propositional symbols. The semantics in this case is the standard propositional logic semantics based on truth valuations, i.e. assignments of 0 and 1 to propositional symbols. This semantics leads to one notion of entailment of attribute implications, the standard propositional logic entailment. This semantics is relevant to our paper because the truth valuations involved may be identified with rows of tables with yes-or-no attributes (table entry  $I(x, y)$  equals 1 iff  $y$  is assigned 1). As a consequence, the propositional logic semantics essentially coincides with the semantics based on tables with yes-or-no attributes. In particular, these two semantics have the same entailment relation which we denote by  $\models^{\text{AI}}$ . The other semantics of attribute implications comes from relational databases and is given by interpreting attribute implications  $A \Rightarrow B$  as functional dependencies in relations [Armstrong 1974; Maier 1983]. We thus have two notions of entailment: first,  $A \Rightarrow B$  may follow from a set  $T$  of implications as a propositional logic formula,  $T \models^{\text{AI}} A \Rightarrow B$ ; and second,  $A \Rightarrow B$  may follow from  $T$  as a functional dependence,  $T \models^{\text{FD}} A \Rightarrow B$ . Fagin [1977] proved that

$$T \models^{\text{AI}} A \Rightarrow B \text{ is equivalent to } T \models^{\text{FD}} A \Rightarrow B. \quad (83)$$

Since the semantics based on tables with yes-or-no attributes, and hence the one based on propositional logic, is a particular case of the semantics based on tables with graded attributes developed in this paper, the following question arises: is there a natural extension of Codd’s relational model of data and the notion of functional dependence in this extension for which

a result analogous to (83) holds? As we show below, the answer is positive. Formally, such an extension consists in replacing ordinary relations in Codd's model by L-valued relations. In particular, the domains in the extended model are equipped with L-valued relations, such as similarity relations, replacing the ordinary equality relations, which are implicitly present in Codd's model and which are utilized e.g. in selection and other queries involving match of tuples. Furthermore, relations on relation schemes are replaced by L-valued relations, which means that a degree in  $L$  are assigned to each tuple. Such degree is generally interpreted as a degree to which the tuple matches a query involving the L-valued relations on domains. Therefore, the L-valued relations have in fact the same meaning in the extended model as relations on relational schemes have in the ordinary Codd's model, namely they are understood as results of queries with the provision that base relations considered as results of empty queries. The above described extension is interesting in its own right because, as the thorough examination in [Belohlavek and Vychodil 2014] reveals, when the L-valued relations on domains represent similarities, the extension plays the same role for relational databases that support similarity queries as Codd's model plays for ordinary relational databases.

For brevity, we restrict to a particular case of the above-mentioned extension of Codd's model, which is sufficient for our purpose. Let us assume that for each attribute  $y$  of relation scheme (attribute set)  $Y$ ,  $D_y$  denotes the domain of  $y$  and that each domain  $D_y$  is equipped an L-relation  $R_y$ . That is,  $R_y$  maps the pairs  $\langle d_1, d_2 \rangle \in D_y \times D_y$  to grades  $R_y(d, d_2) \in L$ , interpreted as grades to which  $d_1$  is related to  $d_2$ . A *data table over domains with L-relations* on  $Y$  we mean a finite relation  $\mathcal{D}$  between the domains  $D_y$ , i.e.  $\mathcal{D} \subseteq \prod_{y \in Y} D_y$ .

*Remark 8.1.* (a) If we require that for every  $d_1, d_2 \in D_y$ ,

$$(\text{Ref}) \quad R_y(d_1, d_1) = 1,$$

$$(\text{Sym}) \quad R_y(d_1, d_2) = R_y(d_2, d_1),$$

$R_y$  may naturally be understood as representing similarity, i.e.  $R_y(d_1, d_2)$  may be interpreted as a degree to which  $d_1$  and  $d_2$  are similar. Furthermore, we may assume that instead of being an ordinary relation,  $\mathcal{D}$  is an L-relation, in which case  $\mathcal{D}(t)$  is naturally understood as a degree to which the tuple  $t$  satisfies a similarity query that involves a similarity query. For instance, assume that the query reads “show tuples with value of attribute age similar to 30”. Then if  $t(\text{age}) = 33$  and if  $R_{\text{age}}(30, 33) = 0.9$ , then the result of such query (applied to a base relation) is naturally represented by a table  $\mathcal{D}$  in which  $\mathcal{D}(t) = 0.9$ . This is basically the idea of the model presented in [Belohlavek and Vychodil 2014].

Notice that if  $L$  is the two-element Boolean algebra, i.e.  $L = \{0, 1\}$  and if every  $R_y$  represents equality in that  $R_y(d_1, d_2) = 1$  iff  $d_1 = d_2$ , the above concept may be identified with the ordinary concept of relation on  $Y$  of Codd's model [Codd 1970; Maier 1983]. From this point of view, while the ordinary model supports queries regarding exact match of domain values, the similarity-based extension supports those regarding approximate matches.

(b) If  $R_y$  is reflexive and L-transitive [Belohlavek 2002; Gottwald 2001], i.e. satisfies (Ref) and

$$(\text{Tra}) \quad R_y(d_1, d_2) \otimes R_y(d_2, d_3) \leq R_y(d_1, d_3),$$

then  $R_y$  is naturally interpreted as a graded preference relation [Richardson 1998].

Ordinary attribute implications, when interpreted in data tables of Codd's model, represent functional dependencies in this model. In basically the same way, graded attribute implications may be interpreted in data tables over domains with L-relations and represent a similar type of dependencies. Namely, an ordinary attribute implication  $A \Rightarrow B$  asserts that the same values on attributes in  $A$  imply the same values on attributes in  $B$ . As we show below, when the L-relations represent similarities, a graded attribute implication asserts that similar values on attributes in  $A$  imply similar values on attributes in  $B$ .

The interpretation of graded attribute implications being introduced follows the basic principles of predicate fuzzy logic [Gottwald 2001; Hájek 1998]. Our aim is to define a degree to which a graded implication  $A \Rightarrow B$  is true in a table  $\mathcal{D}$ . First let us define for any two tuples  $t_1, t_2 \in \prod_{y \in Y} D_y$ ,

$$t_1(A) \sim t_2(A) = \bigwedge_{y \in Y} (A(y) \rightarrow R_y(t_1(y), t_2(y))). \quad (84)$$

Note that  $t_1(A) \sim t_2(A)$  is the truth degree of the proposition “for every attribute  $y$  in  $A$ , the values  $t_1(y)$  and  $t_2(y)$  are  $R_y$ -related” (instead of “ $R_y$ -related” one may use “similar” here and below to obtain the meaning of the particular case with similarity relations). The same way

we define  $t_1(B) \sim t_2(B)$ . The degree  $\|A \Rightarrow B\|_{\mathcal{D}}$  to which  $A \Rightarrow B$  is true in  $\mathcal{D}$  is defined by

$$\|A \Rightarrow B\|_{\mathcal{D}} = \bigwedge_{t_1, t_2 \in \mathcal{D}} ((t_1(A) \sim t_2(A))^* \rightarrow (t_1(B) \sim t_2(B))). \quad (85)$$

According to the principles of fuzzy logic,  $\|A \Rightarrow B\|_{\mathcal{D}}$  is the truth degree of the proposition “for every two tuples  $t_1, t_2 \in X$ : if it is (very) true that  $t_1$  and  $t_2$  have  $R_y$ -related (e.g. similar) values on attributes from  $A$  then  $t_1$  and  $t_2$  have  $R_y$ -related (similar) values on attributes from  $B$ ”.

*Remark 8.2.* (a) One may easily observe that if  $L = \{0, 1\}$  and if every  $R_y$  represents identity, (85) becomes the definition of validity of ordinary functional dependencies in ordinary relations. Furthermore, if every  $R_y$  is reflexive and transitive, and thus represents a preference, we obtain the definition of validity of ordinal dependencies [Ganter and Wille 1999].

(b) The hedge  $*$  in (85) has a similar role as in (37). In particular, if  $*$  is the globalization, see (33), then if  $R_y$ s represent similarities, an implication such as  $\{a_1/y_1, \dots, a_p/y_p\} \Rightarrow \{b_1/z_1, \dots, b_q/z_q\}$ , is fully true, i.e. true to degree 1, in  $\mathcal{D}$  iff similarity to degrees  $a_i$  or higher on attributes  $y_i$  implies similarity to degrees  $b_i$  or higher on attributes  $z_i$ , as mentioned in Section 1. For more information we refer again to [Belohlavek and Vychodil 2014].

(c) In the literature, several approaches to a relational model over domains with similarities and the corresponding functional dependencies have been proposed, [Raju and Majumdar 1988] being among the first ones. As a rule, these approaches lack a clear connection to an underlying logic calculus such as the predicate logic in case of the ordinary Codd’s model or predicate fuzzy logic as in our case. For an overview and comparison of these approaches, we refer the reader to [Belohlavek and Vychodil 2011].

In the rest of this section, we denote by  $\|A \Rightarrow B\|_T^{\text{AI}}$  the degree to which the graded attribute implication  $A \Rightarrow B$  follows from a fuzzy set  $T$  of graded implications in the semantics given by tables with graded attributes, as defined by (43). In much the same way, we define the degree of entailment  $\|A \Rightarrow B\|_T^{\text{FD}}$  in which implications are conceived as functional dependencies in data tables over domains with L-relations:

$$\|A \Rightarrow B\|_T^{\text{FD}} = \bigwedge_{M \in \text{Mod}^{\text{FD}}(T)} \|A \Rightarrow B\|_M \quad (86)$$

where

$$\text{Mod}^{\text{FD}}(T) = \{\mathcal{D} \mid \text{for each } A, B \in \mathbf{L}^Y : T(A \Rightarrow B) \leq \|A \Rightarrow B\|_{\mathcal{D}}\}.$$

denotes the set of models of  $T$ , i.e. data tables in which each  $A \Rightarrow B$  holds to a degree larger than or equal to the degree prescribed by the theory  $T$ .

To answer the question about the relationship between the two concepts of entailment, we need the next two lemmata. Let us define for a given  $\langle X, Y, I \rangle$  a data table  $\mathcal{D}_{\langle X, Y, I \rangle}$  as follows:

- for each  $y \in Y$ , let  $D_y = X \cup X'$  where  $X' = \{x' \mid x \in X\}$  (i.e.,  $X \cap X' = \emptyset$  and  $|X| = |X'|$ );
- for  $x_1, x_2 \in D_y$ , let

$$R_y(x_1, x_2) = \begin{cases} 1 & \text{for } x_1 = x_2, \\ I(z_1, y) \wedge I(z_2, y) & \text{for } x_1 \neq x_2, x_i = z_i^{(\prime)} \text{ for } z_i \in X \ (i = 1, 2), \end{cases}$$

where  $x_i = z_i^{(\prime)}$  means that  $x_i$  is  $z_i$  or  $z'_i$ ;

- $\mathcal{D} = \{t_x \mid x \in X \cup X'\}$  where  $t_x$  is the tuple in  $\prod_{y \in Y} D_y$  for which  $t_x(y) = x$  for every  $y \in Y$ .

As the following lemma shows, degrees of validity in  $\langle X, Y, I \rangle$  coincide with those in  $\mathcal{D}_{\langle X, Y, I \rangle}$ .

**LEMMA 8.3.** *For every data table  $\langle X, Y, I \rangle$  with graded attributes and any graded attribute implication  $A \Rightarrow B$ ,*

$$\|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \|A \Rightarrow B\|_{\mathcal{D}_{\langle X, Y, I \rangle}}. \quad (87)$$

**PROOF.** Let us first observe that

- (a)  $(a_1^* \rightarrow b_1) \wedge (a_2^* \rightarrow b_2) \leq (a_1 \wedge a_2)^* \rightarrow (b_1 \wedge b_2)$  for any  $a_1, a_2, b_1, b_2 \in L$ ;
- (b)  $t_{x_1}(C) \sim t_{x_2}(C) = S(C, I_{z_1}) \wedge S(C, I_{z_2})$   
for any  $C \in \mathbf{L}^Y$  and any  $x_1 \neq x_2$  such that  $x_1 = z_1^{(\prime)}, x_2 = z_2^{(\prime)}$  for some  $z_1, z_2 \in X$ ;
- (c)  $t_x(C) \sim t_x(C) = 1$  for  $x \in X \cup X'$ .

Indeed, due to adjointness, (a) is equivalent to  $(a_1 \wedge a_2)^* \otimes ((a_1^* \rightarrow b_1) \wedge (a_2^* \rightarrow b_2)) \leq b_1 \wedge b_2$  which holds iff  $(a_1 \wedge a_2)^* \otimes ((a_1^* \rightarrow b_1) \wedge (a_2^* \rightarrow b_2)) \leq b_1$  and  $\leq b_2$ . Both inequalities are true. Namely,  $(a_1 \wedge a_2)^* \otimes ((a_1^* \rightarrow b_1) \wedge (a_2^* \rightarrow b_2)) \leq a_1^* \otimes (a_1^* \rightarrow b_1) \leq b_1$  and similarly for  $b_2$ .

Since  $I_{z_i}(y) = I(z_i, y)$ , we have

$$\begin{aligned} t_{x_1}(C) \sim t_{x_2}(C) &= \bigwedge_{y \in Y} (C(y) \rightarrow R_y(t_{x_1}(y), t_{x_2}(y))) = \\ &= \bigwedge_{y \in Y} (C(y) \rightarrow R_y(x_1, x_2)) = \bigwedge_{y \in Y} (C(y) \rightarrow (I(z_1, y) \wedge I(z_2, y))) = \\ &= \bigwedge_{y \in Y} (C(y) \rightarrow I(z_1, y)) \wedge \bigwedge_{y \in Y} (C(y) \rightarrow I(z_2, y)) = S(C, I_{z_1}) \wedge S(C, I_{z_2}), \end{aligned}$$

establishing (b). (c) is evident.

Let for brevity  $\mathcal{D} = \mathcal{D}_{\langle X, Y, I \rangle}$ . We obtain

$$\begin{aligned} \|A \Rightarrow B\|_{\mathcal{D}} &= \bigwedge_{t_1, t_2 \in \mathcal{D}} ((t_1(A) \sim t_2(A))^* \rightarrow (t_1(B) \sim t_2(B))) = \\ &= \bigwedge_{x_1, x_2 \in X \cup X'} ((t_{x_1}(A) \sim t_{x_2}(A))^* \rightarrow (t_{x_1}(B) \sim t_{x_2}(B))) = \alpha \wedge \beta \wedge \gamma, \end{aligned}$$

where

$$\alpha = \bigwedge_{x_1, x_2 \in X \cup X', x_1 = x_2} ((t_{x_1}(A) \sim t_{x_2}(A))^* \rightarrow (t_{x_1}(B) \sim t_{x_2}(B))) = 1$$

on account of (c),

$$\begin{aligned} \beta &= \bigwedge_{x_1, x_2 \in X \cup X', \{x_1, x_2\} = \{z, z'\}} ((t_{x_1}(A) \sim t_{x_2}(A))^* \rightarrow (t_{x_1}(B) \sim t_{x_2}(B))) = \\ &= \bigwedge_{z \in X} ((S(A, I_z) \wedge S(A, I_z))^* \rightarrow (S(B, I_z) \wedge S(B, I_z))) = \\ &= \bigwedge_{z \in X} (S(A, I_z)^* \rightarrow S(B, I_z)) = \|A \Rightarrow B\|_{\langle X, Y, I \rangle} \end{aligned}$$

on account of (b), and

$$\begin{aligned} \gamma &= \bigwedge_{\{x_1, x_2\} = \{z_1', z_2'\}, z_1 \neq z_2} ((t_{x_1}(A) \sim t_{x_2}(A))^* \rightarrow (t_{x_1}(B) \sim t_{x_2}(B))) = \\ &= \bigwedge_{\{x_1, x_2\} = \{z_1', z_2'\}, z_1 \neq z_2} ((S(A, I_{z_1}) \wedge S(A, I_{z_2}))^* \rightarrow (S(B, I_{z_1}) \wedge S(B, I_{z_2}))) \geq \\ &= \bigwedge_{\{x_1, x_2\} = \{z_1', z_2'\}, z_1 \neq z_2} (([S(A, I_{z_1})^* \rightarrow S(B, I_{z_1})] \wedge [S(A, I_{z_2})^* \rightarrow S(B, I_{z_2})])) = \\ &= \bigwedge_{z \in X} (S(A, I_z)^* \rightarrow S(B, I_z)) = \beta \end{aligned}$$

on account of (b) and (a). Therefore,

$$\|A \Rightarrow B\|_{\mathcal{D}_{\langle X, Y, I \rangle}} = \beta = \|A \Rightarrow B\|_{\langle X, Y, I \rangle},$$

completing the proof.  $\square$

Conversely, for a given table over domains with L-relations  $\mathcal{D}$ , define a table  $\langle X, Y, I \rangle_{\mathcal{D}}$  as follows:

- $X = \mathcal{D} \times \mathcal{D}$ ;
- for  $\langle t_1, t_2 \rangle \in X$  and  $y \in Y$ , let  $I(\langle t_1, t_2 \rangle, y) = R_y(t_1(y), t_2(y))$ .

As in the previous case,  $\mathcal{D}$  and  $\langle X, Y, I \rangle_{\mathcal{D}}$  yield the same truth degrees of attribute implications:

**LEMMA 8.4.** *For every ranked data table  $\mathcal{D}$  and any graded attribute implication  $A \Rightarrow B$ ,*

$$\|A \Rightarrow B\|_{\mathcal{D}} = \|A \Rightarrow B\|_{\langle X, Y, I \rangle_{\mathcal{D}}}. \quad (88)$$

**PROOF.** Notice first that

$$\begin{aligned} (t_1(A) \sim t_2(A)) &= \bigwedge_{y \in Y} (A(y) \rightarrow R_y(t_1(y), t_2(y))) = \\ &= \bigwedge_{y \in Y} (A(y) \rightarrow I(\langle t_1, t_2 \rangle, y)) = S(A, I_{\langle t_1, t_2 \rangle}), \end{aligned}$$

and the same for  $B$ . We therefore get

$$\begin{aligned} \|A \Rightarrow B\|_{\mathcal{D}} &= \bigwedge_{t_1, t_2 \in \mathcal{D}} ([t_1(A) \sim t_2(A)]^* \rightarrow [t_1(B) \sim t_2(B)]) = \\ &= S(A, I_{\langle t_1, t_2 \rangle})^* \rightarrow S(B, I_{\langle t_1, t_2 \rangle}) = \|A \Rightarrow B\|_{\langle X, Y, I \rangle_{\mathcal{D}}}. \end{aligned}$$

$\square$

The following theorem answers the question from the beginning of this section.

**THEOREM 8.5.** *For every fuzzy set  $T$  of graded attribute implications and every graded attribute implication  $A \Rightarrow B$  we have*

$$\|A \Rightarrow B\|_T^{\text{FD}} = \|A \Rightarrow B\|_T^{\text{AI}}. \quad (89)$$

**PROOF.** We need to prove  $\|A \Rightarrow B\|_T^{\text{FD}} \leq \|A \Rightarrow B\|_T^{\text{AI}}$  and  $\|A \Rightarrow B\|_T^{\text{FD}} \geq \|A \Rightarrow B\|_T^{\text{AI}}$ . To check the first inequality, it is enough to show that for each  $M \in \text{Mod}(T)$  there is  $\mathcal{D} \in \text{Mod}^{\text{FD}}(T)$  such that  $\|A \Rightarrow B\|_M = \|A \Rightarrow B\|_{\mathcal{D}}$ . This follows directly from Lemma 8.3 by taking  $\mathcal{D} = \mathcal{D}_{\langle X, Y, I \rangle}$ , where  $\langle X, Y, I \rangle$  is a one-row data table corresponding to  $M$ , i.e. with  $X = \{x\}$  and  $I(x, y) = M(y)$  for each  $y \in Y$ . Namely, we then have  $\|A \Rightarrow B\|_M = \|A \Rightarrow B\|_{\langle X, Y, I \rangle} = \|A \Rightarrow B\|_{\mathcal{D}}$ . The second inequality is proved in a similar manner using Lemma 8.4.  $\square$

*Remark 8.6.* The tables  $\langle X, Y, I \rangle_{\mathcal{D}}$  and  $\mathcal{D}_{\langle X, Y, I \rangle}$  constructed from  $\mathcal{D}$  and  $\langle X, Y, I \rangle$ , respectively, are not minimal in size. We use them because their definitions are relatively simple and they do their work in the proofs of Lemma 8.3 and Lemma 8.4.

## 9. CONCLUSIONS

We presented an approach to attribute dependencies for data with grades, such as a grade to which an object is red or a grade to which two objects are similar. Such dependencies extend classical dependencies in Boolean data and classical functional dependencies. We presented results regarding major issues traditionally investigated for such dependencies, including entailment, redundancy and bases of dependencies, associated closure structures, Armstrong-like axiomatization, and computation issues. In addition, we examined a relationship between the new kind of dependencies and the classical ones and showed that the well-known correspondence between attribute dependencies in Boolean data on one hand and functional dependencies in relational model of data on the other hand is retained in the setting with grades but obtains a nontrivial, interesting form. Namely, in the setting with grades, the role of functional dependencies is played by their analogue in an extended relational model in which every domain is equipped with a similarity relation, or a more general binary relation, assigning grades of similarity to pairs of domain elements.

In addition, the paper attempts to make a methodological point, the ramifications of which we consider equally important as the results mentioned in the above paragraph. The point is the following. Classical dependencies are based in classical logic in that the truth value *true* (1) represents presence of an attribute and match of attribute values, while *false* (0) represents absence and mismatch. Moreover, the truth values are manipulated by classical logic connectives and further notions such as validity and entailment of dependencies are derived from classical logic notions. Broadly speaking, classical dependencies are founded in the agenda of classical logic. The presence of possibly many grades in the new situation and the ordinal nature of grades makes the situation challenging and prone to ad hoc treatments, involving for instance metrics representing similarities. Thus, one might attempt to retain the agenda of classical logic, extend the formalism of classical dependencies by a metric (distance function) to represent similarity, and arrive at a blend of a logic-based formalism and a metric-based one. Instead, our approach—like the classical one—is purely logically based, yet capable of handling grades and their semantics in a reasonable way. We consider the grades as truth values in the sense of fuzzy logic, i.e. consider them as truth degrees with 1 and 0 representing the boundary cases and the other ones, such as 0.8, as representing intermediary cases. In a sense, we move to a more general framework, a logical calculus in which statements such as “attribute  $y$  applies to object  $x$ ” and “objects  $x_1$  and  $x_2$  are similar (equal)” are no longer considered bivalent. Rather, these statements are allowed to be assigned, in addition to 0 and 1, an intermediary *grade*, i.e. a *truth degree* between 0 and 1. Such move can effectively be realized. Namely, we argue that data involving grades and reasoning about such data can be modeled utilizing a framework of mathematical fuzzy logic, a recently developed many-valued logic with now well-developed agenda and that this logic may assume the role classical logic plays in the established theories of data dependencies and reasoning about data in general. The main advantage of this approach is conceptual clarity. On the level of syntax, the key notions in the model with grades have essentially the same form as in the classical, bivalent case. This means that the informal description of the key notions in natural language, and hence the intuitive meaning of the key notions, remains essentially the same as in the classical model. Yet, on the level of semantics, grades obtain a proper treatment and permeate the subsequent notions such as validity or entailment in a natural way. Thus, for instance, validity or entailment of dependencies are no longer bivalent concepts. Rather, they naturally emerge as graded notions. One obtains a degree of validity or degree of entailment



of dependencies, corresponding to the idea that a compound statement (such as a dependency claim) involving partially true constituent statements (such as “attribute  $y$  applies to object  $x$ ”) may itself be only partially true, i.e. true to an intermediary degree.

To sum up, utilization of mathematical fuzzy logic as a formal framework for modeling data with grades brings conceptual clarity and makes possible a treatment of attribute dependencies essentially the same way as utilization of classical logic does for data with no intermediary grades. Clearly, the presented approach is not restricted to the problems dealt with in the present paper. In this respect, our paper demonstrates that fuzzy logic is a convenient framework for modeling certain problems that surpass the domain traditionally accounted for by classical logic, namely those that may be characterized by a graded nature of data and reasoning about such data. Such problems abound particularly in situations where human judgment is involved, for which the usage of graded, “fuzzy” notions, such as *red*, *tall*, *similar*, rather than bivalent ones, is characteristic. A further development of theories and methods inspired by such problems presents a challenging and important research goal. The associated research agenda includes several complex issues, some of which we intentionally disregarded in the present paper. One such issue is connected to the fact that the theory we present is not restricted to a particular set of grades and particular (truth functions of) logical connectives on this set. Rather, we proceed in a general way and only assume that the set  $L$  of grades and the logical connectives on  $L$  satisfy certain logically reasonable conditions such as the isotony of conjunction, its commutativity, associativity, and the like. In a sense, the presented theory is qualitative and open to determination of a quantitative component. Clearly, the choice of this component, i.e. a particular set  $L$  and particular connectives on  $L$ , is a step one needs to make when applying the theory. One option in making this step is to proceed on intuitive grounds, which is often the case in applications of fuzzy logic. In fact, there is an argument for considering such option sufficient for practical purpose, namely, that the common qualitative properties of all the potential sets of logical connectives are specific enough to the extent that all the sets of connectives can be considered reasonable for practical purpose. Still, such option may arguably be regarded as too much ad hoc. In fact, the choice of a scale of grades and logical connectives for this scale is a matter that calls for a thorough examination from the point of view of a mathematical and cognitive psychology. In our view, such examination presents challenging problems with broad ramifications and is very much needed.

## ACKNOWLEDGMENTS

Dedicated to Professor Petr Hájek.

## REFERENCES

- Agrawal R., Imielinski T., and Swami A. 1993. Mining association rules between sets of items in large databases. In *Proceedings of SIGMOD*. 207–216.
- Armstrong W. W. 1974. Dependency structures in data base relationships. In *Proceedings of IFIP*. 580–583.
- Belohlavek R. 1999. Fuzzy Galois connections. *Math. Logic Quarterly* 45, 4, 497–504.
- Belohlavek R. 2002. *Fuzzy Relational Systems: Foundations and Principles*. Kluwer.
- Belohlavek R. 2004. Concept lattices and order in fuzzy logic. *Ann. Pure Appl. Logic* 128, 277–298.
- Belohlavek R., Funiokova T., and Vychodil V. 2005. Fuzzy closure operators with truth stressers. *Logic J. IGPL* 13, 5, 503–513.
- Belohlavek R. and Vychodil V. 2011. Codd’s relational model from the point of view of fuzzy logic. *J. Logic and Computation* 21, 5, 851–862.
- Belohlavek R. and Vychodil V. 2012. Formal concept analysis and linguistic hedges. *Int. J. General Systems* 41, 5, 503–532.
- Belohlavek R. and Vychodil V.: Relational Similarity-Based Databases, Part 1: Foundations and Query Systems, Part 2: Dependencies in Data. Submitted.
- Cintula P., Hájek P., and Noguera C. 2011. *Handbook of Mathematical Fuzzy Logic, vol. I, II*. College Publ.
- Carpineto C. and Romano G. 2004. *Concept Data Analysis: Theory and Applications*. J. Wiley.
- Codd E. F. 1970. A relational model of data for large shared data banks. *Communications of the ACM* 13, 6, 377–387.
- Davey B. A., Priestly H. A.: *Introduction to Lattices and Order*. Cambridge University Press, 2002.
- Delobel C. and Casey R. G. 1973. Decomposition of a data base and the theory of Boolean switching functions. *IBM J. Research and Development* 17, 374–386.
- Fagin R. 1977. Functional dependencies in a relational database and propositional logic. *IBM J. Research and Development* 21, 6, 543–544.
- Fagin R. 1999. Combining fuzzy information from multiple systems. *J. Comput. Syst. Sci.* 58, 1 83–99.
- Fagin R. 2002. Combining fuzzy information: an overview. *SIGMOD Record* 31, 2, 109–118.
- Galatos N., Jipsen P., Kowalski T., and Ono H. 2007. *Residuated Lattices: An Algebraic Glimpse at Substructural Logics*. Elsevier.
- Ganter B. 1984. Two basic algorithms in concept analysis. *Preprint 831, Fachbereich Mathematik*, TU Darmstadt.
- Ganter B. 1998. Begriffe und Implikationen (in German). Unpublished manuscript.
- Ganter B. and Wille R. 1999. *Formal Concept Analysis. Mathematical Foundations*. Springer.

- Gerla G. 2001. *Fuzzy Logic. Mathematical Tools for Approximate Reasoning*. Kluwer.
- Goguen J. A. 1967. L-fuzzy sets. *J. Math. Anal. Appl.* 18, 145–174.
- Goguen J. A. 1968–9. The logic of inexact concepts. *Synthese* 18, 325–373.
- Gottwald S. 2001. *A Treatise on Many-Valued Logic*. Research Studies Press.
- Gottwald S. 2008. Mathematical fuzzy logics. *Bull. Symbolic Logic* 14, 210–239.
- Guigues J.-L. and Duquenne V. 1986. Familles minimales d'implications informatives resultant d'un tableau de données binaires (in French). *Math. Sci. Humaines* 95, 5–18.
- Hájek P. 1998. *Metamathematics of Fuzzy Logic*. Kluwer.
- Hájek P. 2001. On very true. *Fuzzy Sets and Systems* 124, 329–333.
- Hájek P. 2006. What is mathematical fuzzy logic. *Fuzzy Sets and Systems* 157, 257–603.
- Hájek P. and Havránek T. 1978. *Mechanizing Hypotheses Formation. Mathematical Foundations for a General Theory*. Springer.
- Hájek P., Holeňa M., and Rauch J. 2010. The GUHA method and its meaning for data mining. *J. Comput. Syst. Sci.* 76, 1, 34–48.
- Hipp J., Güntzer U., and Nakhaeizadeh G. 2000. Algorithms for association rule mining—A general survey and comparison. *SIGKDD Explorations* 2, 2, 1–58.
- Johnson D. S., Yannakakis M., and Papadimitrou C. H. 1988. On generating all maximal independent sets. *Inf. Processing Letters* 15, 129–133.
- Kuznetsov S. O. and Obiedkov S. A. 2008. Some decision and counting problems of the Duquenne-Guigues basis of implications. *Discrete Applied Mathematics* 156, 11, 1994–2003.
- Maier D. 1983. *The Theory of Relational Databases*. Computer Science Press.
- Mayor G. and Torrens J. 2005. Triangular norms on discrete settings. In Klement E. P., Mesiar R. (eds.) *Logical, Algebraic, Analytic and Probabilistic Aspects of Triangular Norms*. Elsevier. 189–230.
- Novák V., Perfilieva I., and Močkoř J. 1999. *Mathematical Principles of Fuzzy Logic*. Kluwer.
- Obiedkov S. and Duquenne V. 2007. Attribute-incremental construction of the canonical implication basis. *Annals of Mathematics and Artificial Intelligence* 49, 77–99.
- Ore O. 1944. Galois connections. *Trans. AMS* 55, 493–513.
- Pavelka J. 1979. On fuzzy logic I, II, III. *Z. Math. Logik Grundlagen Math.* 25, 45–52, 119–134, 447–464.
- Raju K. V. S. V. N. and Majumdar A. K. 1988. Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems. *ACM Trans. Database Systems* 13, 2, 129–166.
- Richardson G. 1998. The structure of fuzzy preferences: social choice implications. *Social Choice and Welfare* 15, 359–369.
- Takeuti G. and Titani S. 1987. Globalization of intuitionistic set theory. *Annals of Pure and Applied Logic* 33, 195–211.
- Tan, P.-N., Steinbach M., and Kumar, V. 2005. *Introduction to Data Mining*. Addison-Wesley.
- Ward M. and Dilworth R. P. 1939. Residuated lattices. *Trans. AMS* 45, 335–354.
- Zadeh L. A. 1965. Fuzzy sets. *Information and Control* 12, 94–102.
- Zadeh L. A. 1988. Fuzzy Logic. *IEEE Computer* 21, 4, 83–93.
- Zadeh L. A. 1994. Fuzzy Logic, neural networks, and soft computing. *Commun. ACM* 37, 3, 77–84.
- Zadeh L. A. 2008. Is there a need for fuzzy logic? *Information Sciences* 178, 13, 2751–2779.